Conceptual Progression in the Domain of Naïve Biology

This thesis is submitted to the University College London Institute of Education for the degree of Doctor of Philosophy by

Zayba Ghazali-Mohammed

DECLARATION OF ORIGINALITY:

I, Zayba Ghazali, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

DECLARATION OF WORD COUNT:

Word count (exclusive of appendices and bibliography): 89,846 words. This is in accordance with the institute word limit of 100,000 words. Approval for an extension was granted by the Director of Postgraduate research.

ACKNOWLEDGEMENTS:

Undertaking this PhD has been the most incredible journey, with tremendous highs and equally tremendous lows. In last six months of writing, I have come to realise that there are several people who deserve a mention of gratitude for their on-going support, guidance, and wisdom throughout these past three years.

Firstly, a sincere thank you goes to my supervisor, Andy Tolmie. Not only has his enthusiasm and keen interest in my work been encouraging, his huge support, wit, and patience has been calming in these frequent stormy waters known as research. Having a supervisor that has many of the same perfectionist characteristics in their work as I do has also been reassuring!

I am also immensely grateful to my friends and family for two things in particular: proofreading my work, of course, but more importantly providing me with a sense of reality. Their repeated questioning about my research and (perhaps feigned?) interest encouraged me to go on and pulled me out solitary confinement when I really needed it. Without them the past three years would not have been possible.

Gratitude is also due to the teachers and children who took part in my study. The many hours spent testing in schools have been very rewarding and I hope the outcomes of this thesis may one day return the huge favour they granted me.

Finally, this research was made possible by the doctoral scholarship I was awarded by the Economic and Social Research Council (Grant number: 1238660).

ABSTRACT:

This study aims to explore children's conceptual development in science. Curricular design is an activity that rests on some fundamental assumptions about the organisation of knowledge and development of understanding. Presently the National Curriculum for England is organised in a manner that assumes sequential learning of scientific concepts: generalised understanding can be developed on the basis of established concepts. However there is a distinct lack of any systematic literature on the processes behind conceptual development. What little research there has been has often shown children's learning to be piecemeal (e.g. diSessa, 1993). Therefore questions into how the mind and cognitive schemata are organised cannot be fully answered without first exploring the potential processes behind conceptual change and above all the ways in which related concepts are coordinated and interlinked, something that has rarely been the focus of psychological investigation. The present study explores young children's (aged 4-11) ideas about biological phenomena in an effort to understand conceptual development using a triple-cohort longitudinal design. Children were recruited into 3 cohorts spanning the primary age-range and followed-up a year later so that a putative developmental trajectory relating to the understanding of biological concepts could be ascertained. Children were assessed using general cognitive and demographic measures. A novel interview method was developed to assess children's biological knowledge. Findings revealed children's understanding to be heavily influenced by biologically-specific language, which may act as a mechanism for conceptual change. It was also found that general cognitive abilities and demographic factors had very little influence on conceptual change in the domain of naïve biology, in contrast to previous research. Finally, children's knowledge appeared to be more theoretical than hypothesised, as related biological concepts were predictive of each other. Given these findings, a new model of conceptual organisation and change is proposed.

TABLE OF CONTENTS

For my parents

"Try to learn something about everything, and everything about something."

-Charles Darwin

INTRODUCTION

CHAPTER 1

1.1 The current picture of science

Trying to understand how children view the world has intrigued curious minds for decades. While there has been much progress in our collective understanding of how children acquire concepts, there is plenty yet to uncover. These endeavours are important. Towards the 1980s there was a sudden interest in children's naïve understanding about the world and specific questions were asked about what the content of children's concepts might be at different ages, how these concepts might develop with age, and how they might be organised. Indeed the answers to these questions are highly relevant, not just in terms of theoretical research but also in terms of the educational implications involved. By knowing when a child is cognitively equipped to learn a concept, the mechanism(s) the child employs in order to learn it, and how these concepts may go on to influence the development of each other into more sophisticated ideas, it is possible to shape teaching methods and curricula around this knowledge, and enhance education as a practice.

Learning science requires children to *think scientifically*. This term is usually taken to mean two things: firstly, the actual process of scientific thinking e.g. examining evidence and problem solving. Secondly, the understanding of phenomena and the mechanisms involved. The second definition of scientific thinking encompasses the first and focuses mainly on the conceptual understanding and integration of scientific concepts. This is the area that the thesis will specifically focus on.

Unfortunately, science is an area that many children struggle to learn and a failure to do so can lead to a decline in the number of individuals taking up science related degrees in higher education, to the detriment of the national economy (Royal Society, 2010, 2011). The failure of many children to adequately learn science has also had a negative impact on scientific literacy in the broader population, limiting the capability of people to participate meaningfully in debate and decision-making about policy on science and scientific developments.

The impact of poor science learning is less well quantified in comparison to the impact of poor literacy or mathematics learning. Poor performance in science, literacy, and numeracy has been shown to affect economic growth and income (Wilson, 2009). For example, Beddington (2008) showed that a failure to address learning difficulties in literacy and numeracy leads to poor educational attainment and employment prospects. Bynner and Parsons (2005) have shown that in particular, poor numeracy skills are associated with higher risk of unemployment, depression, crime, and ill health. Thus, it is evident that literacy, numeracy, and science are core areas with which to focus study on, not least in the school curriculum, where these three areas have remained the key focus of comparative international assessments (Tolmie, 2012).

Given the huge impact these topics have economically, developing effective pedagogies in each of these areas (but science in particular given the lack of research) should be a priority and not based on arbitrary principles or political agendas, but rather on sound scientific research (Tolmie, 2012; Ghazali, 2015).

Consequently, it becomes necessary to investigate the range of understanding children have about certain scientific concepts and how this knowledge might progress i.e. conceptual development of scientific concepts. By doing so, researchers will gain a better understanding as to why some children have difficulties with science from an early age, and investigate whether curricular structure, content, and pedagogy itself could be enhanced to promote better grasp of scientific concepts.

The difficulty with this is that conceptual development is an ill-defined concept, which has left a lack of consensual agreement amongst researchers about its nature. Many researchers have not been overtly clear about what they mean regarding concept development or change. Is it the case that all types of concepts change in the same way so that as more knowledge is obtained, concepts change using the same system and continue to do so? The answer to this question has reached very little consensus among researchers, specifically about how initial ideas might develop into more sophisticated concepts. If the answer is yes, this would imply that conceptual development and conceptual change are the same, when it is not yet clear whether change involves restructuring existing knowledge, and development involves encompassing new concepts into existing knowledge. It may even be that the way in which initial concepts are organised goes onto constrain future acquisition and organisation of new concepts. Consequently, a failure to address the nature of change means that there is not yet adequate evidence to inform theories about the mechanisms involved.

Due to these limitations, a definition of conceptual progression is needed. In the past it has simply been discussed as the journey between a naïve and rudimentary (perhaps incomplete or even inaccurate) understanding of a concept, to a more sophisticated and expert-like understanding of the same concept (Karmiloff-Smith, 1988). For example, knowing that “plants need sunlight to grow” to the more sophisticated causal/mechanistic understanding of “plants need sunlight to allow photosynthesis to occur, enabling the plant to function and consequently grow.” As of yet, there have been few attempts to develop a principled scheme to measure conceptual change and without this it becomes difficult to judge conceptual progression, developmental trends of this progression, and its influencers.

The problem is further accentuated by the fact that what little cross-conceptual (conceptual development and progression across multiple concepts) research there has been, is suggestive of piecemeal development (e.g. Tolmie, Tenenbaum, & Pino-Pasternak, 2009; Gelman & Markman, 1986) and therefore conceptual understanding is unlikely to progress in such a straight-forward and linear fashion, contrary to earlier theorising (e.g. Piaget, 1972). Other work also highlights the fact that young children start school with existing prior knowledge and often hold misconceptions that are highly resistant to change (Driver, Guesne & Tiberghien, 1985). The fact that children demonstrate existing conceptual knowledge prior to any formal instruction, is something not accounted for by the current model of teaching in the United Kingdom (UK).

1.2 The National Curriculum

Curricular design is an activity that rests on some fundamental assumptions about the organisation of knowledge and development of understanding. Presently the National Curriculum (NC) for England and Wales is organised in a manner that assumes sequential learning of scientific concepts so that generalised understanding can be developed on the basis of earlier concepts. For instance in Key stage 1 (KS1) children are taught the differences between living and non-living things and in Key stage 2 (KS2) they are taught to establish causal effects and explain how living things and non-living things work (Department for Education; DfE; 2001). However, even if it may seem intuitively reasonable, the assumption that seemingly complex concepts can only be understood after acquiring simpler ones might be premature due to an overwhelming lack of research. Thus far, the majority of research in conceptual change in science has tended to focus on the growth of concepts in an isolated fashion, without considering the relationships between them (e.g. Carey, 1985; Keil, 1996). Hence, very little is known about the organisational principles that might be at work, despite the fact that there is scant evidence for children being *tabula rasa* prior to any formal instruction.

Another issue with regards to current curricular organisation is that assessments of academic achievement are also based on these seemingly problematic structures, which have implications for the reliability of current measurements of children's knowledge. Indeed, it may be this very assumption past approaches have taken that has most likely driven curricular design and organisation to be based on little more than taxonomic structure and evident conceptual precedence typically seen from the viewpoint of the expert rather than of the learner, for whom emergent structures of knowledge may be organised in different ways. If so, it is possible that curricular content as it stands may not be in concordance with the nature of conceptual change both within and *between* concepts in children, and it serves to purpose that the curriculum itself be under scrutiny and possible amendment, therefore.

Naturally, political agendas have also contributed to the changes in the current education system in England and Wales. The NC has had the merit of ensuring that some attention to curriculum planning was taking place, at least at primary level, rather than it being driven solely by individual teacher intuition and school resources. However, this has also led to some seemingly unwarranted and disruptive changes. More recently, a draft of the 2014 NC was released in 2011 and there was much uproar and speculation about the bases for drastic changes in curricular topics, after what seemed like a lack of evidence in support of these changes. This thesis will only focus on the new primary science curriculum for KS1 and KS2, introduced from September 2014, and for reasons that will be enlarged below, children's understanding of biological phenomena in particular.

The weaknesses of design approaches towards examining conceptual understanding of scientific phenomena has consequently led to a weak evidence base with which to ground the recent changes in curricular content and organisation. In the context of biology, these changes include the introduction of more practical science and investigation to encourage scientific thinking in terms of the processes involved in experimental research, the introduction of using more scientific terminology from a young age, learning about biographies of notable scientists, and perhaps most unexpectedly, the introduction of evolutionary concepts in primary school that had only ever been previously taught in secondary school.

While basic evolutionary concepts such as variation and inheritance were already on the primary curriculum (DfE, 2001), it was not typically until secondary school where evolution was formally taught as a module. Although children might already be exposed to evolutionary concepts, there is little systematic research in this area to support any views on whether teaching evolution at primary school level is wise. In fact, despite a fair amount of research around the nature of children’s scientific learning, the outcomes of this research have been very fragmented so there is no real basis for informing any wider organisation, or indeed the curricular changes for 2014.

In sum, the current structure of the NC might not be matched up to the ways in which children understand concepts over time. Exactly how concepts are changing, developing, and progressing is currently unknown in any definitive sense. Questions into how the mind and cognitive schemata are organised cannot be fully understood without first exploring the processes behind conceptual change and above all the ways in which related concepts are coordinated and interlinked, something that has rarely been the focus of psychological investigation.

1.3 Conceptual domains

Researchers have put most effort into understanding initial acquisition and development of key areas of learning on the grounds that these are the immediate priority, both theoretically and educationally. For instance, the cognitive changes involved in literacy development are best understood (e.g. Hulme & Snowling, 2009) followed by numeracy and the early stages of arithmetic (e.g. Dowker & Sigley, 2010). Understanding in these areas has been used to aid teaching practice, and yet comparable work has not been undertaken in the realm of science, with researchers frequently disagreeing about the nature of concept formation and concept change. This has led to a seemingly unsystematic approach to investigating conceptual development in science in comparison to the work in literacy and numeracy (Tolmie, 2012). Moreover, the work into literacy and numeracy is unable to shed light onto the nature of conceptual development in science. The initial development in literacy and numeracy is focused to a much greater extent on performance outcomes (e.g. Gathercole, Brown, & Pickering, 2003; Alloway, Gathercole, Kirkwood & Elliot, 2008), and the conceptual dimensions involved are substantially narrower than is the case in science, where any phenomenon requires a distinct conceptual structure. For example in literacy, it involves learning of orthographic/sound relationships, and development of the ability to extract meaning from groupings of text (Hulme & Snowling, 2009). For numeracy, development involves a grasp of counting procedures and principles, recognition and use of symbols, translation between verbal and symbolic formats, and knowledge of number facts and calculation strategies (Soltesz, Szucs, & Szucs, 2010). More fundamentally, the underpinning skills and effective pedagogical strategies involved in literacy and numeracy are both specific and distinct; hence given the further differences between science and these two areas, there is clear implication that science requires separate systematic investigation. The organisation and the development of skills and concepts in various areas is simply too different, and research into the specificity of conceptual development corroborates this (see Hirschfeld & Gelman, 1994).

1.4 Conceptual sub-domains

One potential factor that increases the difficulty of understanding conceptual progression in science is the broad areas within science itself: chemistry, biology, and physics. Each area deals with quite differentiated, complex concepts that are often highly theoretical in nature. Therefore it may be unhelpful to assume that conceptual progression in all scientific areas is likely to be the same because the nature of conceptual growth might differ according to the types of phenomena involved. For instance, understanding in physics can be quite atomistic or piecemeal as there is less logical reason for concepts to be connected. The very nature of physics is to deconstruct broad concepts into their fundamental principles that are often unconnected in character. For example, in state change, there is a basic generalisation for ideas about energy, processes, and exhibited physical state across different types of state change. However, unless one studies advanced physics, there is little extra information a child requires to merge these particular concepts beyond those basic and somewhat isolated ideas, hence a need for deep conceptual integration is less likely.

In contrast, there has always been recognition of the complex systems involved in biology as the very nature of *this* discipline seeks to take smaller concepts and merge them together to understand various phenomena. For example, in a simple concept of plant biology, a child may know a plant requires water, sunlight etc. to grow. Thus, biology is an area that focuses on combinations of interacting processes, whereas physics is primarily concerned with separate, if not additive processes.

For these reasons a focus on biology might be productive, at least initially, in trying to understand the nature of conceptual progression in science. The inherent similarities between different contexts makes it a better area to look for connections between concepts and observe how these influence the growth and development of each other. For example, biological concepts such as inheritance are naturally connected to aspects of biodiversity, ecology, and even evolution; therefore in terms of trying to examine conceptual integration and progression in science, biology is a suitable area to begin investigation because arguably more simple concepts of inheritance may be gradually enlarged and built upon to form a more complex understanding of something that requires prerequisite knowledge of inheritance, such as evolution. Put simply, there is more reason to expect connectedness among biological concepts, and so this makes it a better place to begin observation of conceptual development and change in comparison to other scientific areas. Principles concerning the nature of connections found in biology may not extend to other areas such as physics, but these provide at least an initial steer on what to investigate; a finding of apparent 'unconnectedness' in physics would not serve the same purpose.

1.5 Overview of the thesis

The section above outlined the potential benefits of investigating the area of biology, which will allow examination of how primary school children, working to the NC for England and Wales, are developing related concepts, and how these concepts might change and integrate over time. It will also allow examination of past approaches to conceptual change research that may have ultimately led to the current taxonomic structure of the NC, which in turn will allow closer inspection of the curriculum and areas of potential amendment.

For these reasons, this thesis aimed to examine three main questions with regards to children's biological knowledge, given the current literature:

1) What can be said about the ways in which children are acquiring initial rudimentary and perhaps inaccurate biological concepts (henceforth referred to as naïve biological concepts), and how these develop into more sophisticated ones?
2) How do concepts seem to be structured and organised, and more specifically, what constitutes conceptual development in areas of biology in young children? Can a principled scheme to measure this be empirically developed?
3) Are there any other factors, such as environmental or social factors, besides conceptual progression alone that influence the conceptual development of children's biological ideas, such as general cognitive abilities?

To address these general objectives, Chapters 2-5 explore the theoretical background to the areas outlined above, and lead towards an overview of the more specific research aims.

Chapter 2 outlines research into naïve biology in particular where the general focus in the past has been on natural kinds and inheritance research with a particular emphasis on children's tendency towards essentialism, which the majority of research into naïve biology has been influenced by. Research in other related biological areas to inheritance (biodiversity, ecology, and evolution) is discussed before speculating on a route of conceptual progression. The potential flaws of previous research paradigms are discussed, and alternative plausible explanations are offered to account for past research findings in naïve biology including children's natural predisposition to categorise and observe regularity.

Chapter 3 introduces the notion of conceptual change and ideas of domain-generality and domain-specificity of knowledge. This chapter also introduces the key debates surrounding the basis of conceptual development and organisation of conceptual structures of knowledge; highlighting the view that knowledge is fragmented, versus the view that knowledge is theoretical. Finally, this chapter introduces a possible route towards knowledge acquisition where language and collaborative learning appear to be key drivers of change.

Chapter 4 specifically deals with alternative theories that move away from the potentially flawed essentialist paradigm, starting with statistical learning and categorisation, introduced at the end of Chapter 3. The differences between implicit and explicit thought are highlighted alongside recent insights from neuroscientific studies to promote views of dual-route learning models of conceptual change. The possible importance of general cognitive abilities is also discussed by examining research about their importance for literacy and numeracy, and their potential influences on science learning, particularly the related role of working memory and inhibitory control. Given these insights, the relationship between science and numeracy is also considered before examining the role of these influences within biology specifically, due to their relevance for systems thinking and the potential impact of this on conceptual growth.

Chapter 5 offers a summary and overview of the theoretical background and goes on to provide a general overview of the rationale for the proposed research. The key research questions will also be discussed followed by a timeline of data collection.

Subsequently, reports on two pilot studies (Chapter 6) are outlined to address the issue of defining conceptual advance in an empirical fashion, before the methodology chapter (Chapter 7), which describes the triple-cohort longitudinal design. The longitudinal study spans two years and tracks children's understanding and progression of biological concepts between cohorts across the primary age-range. Therefore the findings will be explored over three results chapters; the first (Chapter 8) outlining findings from the data collected at Time 1, and the second (Chapter 9) outlining findings from the data collected at Time 2. The final results chapter (Chapter 10) attempts to bring the results from the longitudinal study together and discusses statistical models that were constructed in order to map conceptual development of biological knowledge across primary school.

Lastly, the discussion Chapters 11 and 12, begin by exploring key points from Chapter 10 in a theoretical discussion. This is followed by Chapter 12, which summarises findings with regards to educational practise, and explores limitations and future directions.

CHAPTER 2

2.1 Naïve Biology

Chapter 1 highlighted the issues and lack of agreement regarding conceptual organisation and the process of conceptual change. The process of concept progression has only been explored in limited fashion. Many theorists agree that there is indeed a process of change, yet work has typically focused on tracking change in single constructs over a relatively short time frame (because of difficulties with designing measures that have wider applicability) using cross-sectional methods and no assessment of input. Thus what is currently known about conceptual change is somewhat piecemeal and there is no clear consensus about what conceptual change actually involves, although there is agreement that it is the development of a rudimentary concept into a more sophisticated or complex form. There is also agreement on the fact that this development does not always lead to an accurate conception. This is discussed further in Chapter 3.

The lack of consensus within research has consequently led to a variety of ideas about how children might be acquiring biological concepts. Although children might have fairly sophisticated ideas about biological phenomena before any formal instruction, it is not yet clear how their prior concepts are influencing any new information they acquire or what the nature of conceptual organisation and progression in biology might be. In order to discuss what conceptual progression of biological concepts might look like, one must first consider the degree of understanding children aged 4-11 have around aspects of naïve biology. This will now be explored.

One of the ways that conceptual change research has been approached is to define the different domains of knowledge children have, because appropriate models of the organisation of knowledge are required for adequate theorising about conceptual change. It has largely been thought that the ways in which children group together a set of representations sustaining a specific area of knowledge such as language, number, biology etc, is known as a domain (Karmiloff-Smith, 1992). Hirschfeld and Gelman (1994) describe a domain as:

> "…*A body of knowledge that identifies and interprets a class of phenomena assumed to share certain properties and to be of a distinct and general type…functions as a stable response to a set of recurring and complex problems…involving difficult-to-access perceptual, encoding, retrieval and inferential processes dedicated to that solution*" (p.21).

These domain models of conceptual change widely accept the view that human beings are endowed with a set of reasoning abilities or a common set of processes that apply to all thought (Hirschfeld & Gelman, 1994). These are typically known as domain-general models of cognition. However, there is now the increasing view that many of these cognitive abilities are specialised to handle specific types of information such as categorising natural kinds etc, and these are known as domain-specific cognition models (Hirschfeld & Gelman, 1994). Depending on the way children organise and manipulate the information they acquire, knowledge might be theoretical or fragmented. This is still currently an ongoing debate in the literature that is still relatively unresolved.

Part of the problem lies in the fact that learning science depends on more than just acquiring new information, rather it requires conceptual change (Zaitchick et al., 2013; Nersessian, 2003). Simply acquiring facts with age does not necessarily suggest that one is able to hold a theoretical understanding. Much of the previous work within the area of naïve biology has focused on children's early or naive ideas about natural kinds and inheritance, and has been taken to indicate that children could have fairly coherent concepts by age 5 (Springer, 1999). Inheritance describes the passing of traits from parents to offspring and is one of the most studied areas in biology. There have been various theories offered to account for children's understanding of inheritance from around age 4 but by far ideas of essentialism have been most frequent.

2.2 Inheritance and Essentialism

Essentialism is the view that categories have an underlying reality or true nature that one cannot observe directly but that gives an object its identity (Gelman, 2003). Essentialism has been argued to be a potential early precursor of genetic concepts, and refers to the idea that natural kinds contain 'essences' of their being which makes them what they are (Gelman & Kremer, 1991). The essence therefore gives rise to the observable similarities shared by members of a category i.e. children's understanding about inheritance (Gelman, 2003).

Psychological essentialism, according to Gelman (2009, 2015), is the tendency to believe that categories are real, natural, and have a deeper basis and hidden causal essence that may or may not be implicit. Gelman, Coley and Gottfried (1994) conclude that essentialism stems from a domain-general assumption; children assume that events and features are caused and appear biased to search for internal inherent causes.

In fact, many of those who investigate children's understanding of naïve biology argue that children often hold coherent thoughts about biological phenomena and display evidence of theoretical thinking. For instance, Keil (1996) demonstrated that children are able to understand that characteristic features on their own should not be used to diagnose membership in a kind. In one particular study, children were told stories about goats that were given vitamin shots when they were born so that they did not appear to look like goats, but had white curly coats and said "*baa baa*." Children were then asked whether these animals were still goats or sheep (see Keil, 1996). It was found that depending on the mechanisms responsible for the changes, exactly the same types of characteristics were judged either as having no effect on the animal involved or changing it to an entirely new kind. For these reasons, Keil argues that children embed their concepts and interpret their properties in a variety of connected and theoretical beliefs, ultimately leading to a coherent understanding.

It is often the case that essentialist responses from children have been used as evidence of coherent biological reasoning, which in turn have been used to provide evidence for a domain of naïve biology (Carey, 1985; Altran & Medin, 2001; Soloman, 2002). For example, a study by Gelman (2003) examined children's inheritance beliefs about gender using a nature/nurture task. Children were given a story about an infant girl raised only by men (or an infant boy raised only by women) and asked to infer various properties about the boy or girl when s/he was 10 years old. When asked whether the boy would play with a truck or a tea set, a boy aged five replied: "*because boys play with boy things and girls play with girl things*." in comparison to a boy aged ten "*because usually, since she has a girl brain, she'd like to play with a tea set*." (See Gelman, 2003, p.97 for more examples). These examples demonstrate that although the 10-year-old does seem to be alluding to some kind of underlying mechanism for why girls like to play with tea sets, for example it may be argued that all the 5-year-old is doing is asserting an observable probability and no sense of mechanistic understanding is evident here at all and yet the findings from this study were taken as evidence of coherent understanding about inheritance, which seems to be a grossly overreached conclusion. Therefore, despite some early findings indicating that children's inheritance concepts could be theoretical, it was often the case that early methods used to examine this had several issues.

The findings from studies such as Keil's (1996) have generally stemmed from the consistent use of experimental tasks of which, according to Gelman (2015) there are three types: inductive projections, transformation, and switch-at-birth tasks. The first is where children make sensible projections of newly taught biological properties. For example, Carey (1985) demonstrates that when 4-year-old children are taught novel and vague biological information such as "*a spleen is a round, green thing inside people*," children projected the new organ information to other animals with decreasing likelihood from mammals through to birds, insects, and worms. These patterns of induction are constrained by children's existing concepts of *animal* and *living thing*, hence learning of novel biological information is taking place by linking the new concept to one(s) that children already have basic ideas about. The second task, transformation, is the type of task most common in this paradigm. For example, Keil (1996) has shown when asking pre-schoolers whether a scientist could ever perform a new operation to a toy bird to make it into a real bird, even young children thought it was impossible. This, Keil argues, is evidence of a rudimentary understanding of inheritance.

Finally, the last task is the "switch-at birth" vignette. Children hear a story where a child or animal is adopted from its biological family to a host family that looks dissimilar to the child or animal, and children are asked to predict what traits may or may not be inherited by the child e.g. temperament, hair colour, weight (Keil, 1996). Gelman (2015) describes another variation of this task, which suggests that children think internal organs have causal consequences. In this study, children aged 4-7 years were shown heart transplant vignettes and asked "*if you had X's heart inside you, would you be meaner*?" (Note X is previously described as being mean). They found the majority of children said yes in answer to the question and took this as evidence to suggest that children have internal essential beliefs.

All of these paradigms that have been outlined so far rely on gathering insights into children's understanding about biological concepts via inference based on bridging together their responses to hypothetical scenarios, a potentially problematic strategy. However, there is increasing research to suggest that children's apparent reasoning may be little more than a reflection of children's cultural or behavioural knowledge in the absence of any biological causal or mechanistic understanding (Altran, 1987; Soloman, 2002). For example, the language children use during investigations has been taken as evidence of coherent knowledge of inheritance, yet often the criteria used to interpret language as coherent are abstract, and interpretations may therefore be erroneous (e.g. Gelman, 2003).

It may be that a child holds initial rudimentary thoughts that are un-sophisticated and possibly tacit in terms of when inheritance concepts are acquired, for example if children are acquiring perceptions of co-varying regularities in the environment and thus are gaining some implicit knowledge about particular biological phenomena. Hence a young child may have no real sense of inheritance other than the patterns of observed regularity in the environment making it highly probable that organisms give birth to the same kind (perhaps with some degree of individual variation). However past research has all too often focused on children's understanding of interspecies variation when inheritance covers more than just knowing species run true. It covers a whole host of sub-concepts that relate to growth, genetics, and intra-species level variation, which naturally relate to ideas around biodiversity, ecology, and evolution.

This is further accentuated by the large body of work around inheritance where the essence of an organism gives rise to observable similarities shared by members of the same category (Carey, 1985; Gelman, 2003 p.7; Shtulman, 2010; Gelman, 2015), or in other words, allows the child to understand interspecies variation.

However, the very nature of essentialism means children will assume members of a species are alike so by default should *fail* to understand variation of members *within* a species (Shtulman & Shultz, 2008). Indeed it is exactly the idea of *intra*-species diversity that the body of research following the essentialist paradigm has somewhat ignored. In fact, if one is ever to understand how children's concepts are progressing, knowing that species run true is not the helpful part of the concept because it only captures biodiversity at the interspecies level whereas the crucial element is the significance of individual variation within a species.

This also has obvious consequences if we consider basic evolutionary ideas about successive adaptation. A child must be able to grasp that individuals within a species adapt to their environment and this perpetuates over time, otherwise a thorough understanding of evolution will fail to be achieved (Shtulman, 2012). The consequences of this, therefore, are that inheritance concepts may actually be less accurate in a scientific sense than previously thought. The key point to note is the difference in children's conceptual development between inter- and intra-species variations, which has scarcely been investigated if at all. Thus what is needed is a methodology that is sufficiently sensitive to grasp both types of diversity, as this will allow one to investigate routes of developmental progression against other biological phenomena. Shtulman and Schultz (2008) suggests that knowing a species' identity and categorising it allows us to determine where the organisms should live, for instance categorising an organism as a fish, means also knowing that it must live in water, so in this sense categorisation is driving knowledge about habitat and ecology, which may then in-turn constrain the idea further to allow children to subsequently understand certain habitats, such as fresh water, are home to certain types of animals, such as salmon, building up a larger web of connected knowledge related to key initial inheritance concepts (Assaraf & Orion, 2007., Hipkins et al., 2008).

2.2.1 Current work in inheritance

As stated earlier, more recent work is starting to support the idea that children's concepts around inheritance and other biological phenomena might be less theoretical than previously thought, even among older children. In one study (Williams, 2012) children's knowledge about two connected naïve inheritance concepts were explored using a modified version of the phenotypic similarity task (where participants predicted and explained feature outcomes in both an offspring and a sibling). In the second task, participants offered explanations for instances of parent-offspring dissimilarity and grandparent-offspring resemblance (phenotypic difference task). Findings revealed significant age trends between 4-10 years in naïve inheritance concepts. However, there was little consistency in children's inheritance explanations within or across tasks, indicating fragmented ideas on ostensibly related phenomena. This suggests that it is possible for children to hold accurate knowledge, however without understanding how or why things are as they are, they cannot be thinking about biological phenomena in a theoretical manner.

Similarly work by Williams and Smith (2010) found that although children's (aged 4-14) inheritance knowledge became more consistant and accurate over time, children still failed to explain this knowledge, even at 14 years. This study specifically sought to examine the effects various tasks have on children's inheritance knowledge as if children's knowledge was truly coherent, different tasks should not invoke different responses. Three different types of task were used to examine the effect that task had on children's inheritance knowledge. These were: a modified adoption task (children were required to distinguish traits inherited by between biological and social parentage), a causal mechanisms task (investigating children's preference for various mechanisms of inheritance), and a family relatedness task (children were required to provide judgements on the relatedness of family members and explain these judgements). The study found that children were able to demonstrate increasingly good knowledge in kinship judgement and phenotypic similarity tasks, but when given a task requiring children to explain causal mechanisms, there was significant variability in children's understanding with children having either partial, inaccurate, or no knowledge in this area, even at age 14.

The results from this study illustrate how the onset of inheritance concepts can be very complex, as others (Inagaki & Hatano, 2002; Siegal & Peterson, 1999) have also found. It may be that children have early concepts that are implicit (see Medin & Altran, 1999; Wellman & Gelman, 1992) that later undergo conceptual re-organisation and change to become explicit concepts. This will be discussed further in Chapter 3.

The findngs from Williams and Smith (2010) highlighted that younger children's concepts are tied to a particular task where some inheritance concepts were explained via social terms e.g. in the famly relatedness task, and others were explained via biological terms e.g. the adoption and causal mechanisms task. This indicates that younger children appear to treat a variety of tasks differently and employ different aspects of their inheritance knowledge, ultimately suggesting that it has a fragmented structure. This changes somewhat by age 14 where adolescents demonstrated consistent reasoning across all tasks. Williams and Smith (2010) explain these findings by suggesting that adolescents' implicit knowledge has become refined and allows them to recognise similarities across tasks investigating the same underlying biological phenomenon (inheritance through genetics).

This implies that inheritance concepts become more theory-like with age but initially, are very fragmented.

This finding of incoherence and fragmentation is consitent with children's other biological concepts such as life and death (Slaughter, Jaakkola & Carey, 1999) and illness (Myant & Williams, 2005; Williams & Binnie, 2002). Interestingly however, there was not always consistency in adolescents' inheritance explanations, hence although they can demonstrate inheritance knowledge in all tasks, they did not always verbalise this in open-ended questions, perhaps because few children had a detailed grasp of genetics to offer sufficient explanations of inheritance. Thus these findings are again at odds with those who suggest conceptual development is a process of theory change (e.g. Gopnik & Meltzoff, 1997).

As well as examining whether or not children have the same inheritance concepts across different tasks, research from the field of science education has also demonstrated that children have multiple concepts for one scientific phenomenon and that these different concepts might be employed depending on the type of context (Engel, Clough & Driver, 1986). Hence it becomes reasonable to assume that the same may also be true for biological concepts such as inheritance, where chidren's ideas may appear more coherent in particular contexts than in others, resulting in researchers to hold differing opinions about the true nature of children's understanding.

diSessa (1988) argues that children's inheritance concepts may be tied to particular task contexts because they may be contextually rooted in personal experience. Hallden (1999) also supports this idea and suggests that this may lead younger children who have less knowledge of inheritance to use both biological and social concepts of inheritance interchangeably (which is something that Carey (1985) also found in her research) depending on the task used to examine children's knowledge. It follows then that older children who have acquired more explicit inheritance knowledge start to use this knowledge consistently in their reasoning and are less susceptible to changes in context i.e. knowledge begins as fragmented but may become more theoretical with age and increasing experience. This may also lead to more consistent reasoning about inheritance-related phenomena with increasing age too (Williams & Smith, 2010).

The more recent work described above not only aids our understanding about whether children employ different concepts in different task settings, but it helps to answer the wider questions of what the true nature structure of children's knowelge is: theoretical or fragmented. Although past work indicated that children's concepts are theory-like and cohesive in structure (Gopnick & Meltzoff, 1997; Gopnik & Wellman, 1994; Slaughter & Gopnik, 1996) recent work has highlighted the constraining aspects of judgements tasks that do not allow for a full examination of children's reasoning and understanding. Instead, recent work that has tried to consider this illustrates that children's concepts are fragmented and relatively uncoordinated in structure initially (Williams, 2012; Williams & Smith, 2010; Karmiloff-Smith, 1992; diSessa, 1988). At what point children's knowledge starts to become theoretical is therefore still unanswered.

However, something that is still relatively unexplored in the literature is the effect theoretical or fragmented inheritance knowledge might have on other related biological concepts. This is important to investigate because if knowledge is theoretical as has been argued in the past then children should be able to use this knowledge to also explain other biological concepts. However if knowledge is more fragmented as recent work is beginning to show, it is likely that children's knowledge is piecemeal across all platforms and therefore investigating how and when integrated knowledge comes about is of great importance for educators.

2.2.2 Related biological constructs:

Assessing naïve biology requires careful consideration given the fact that past research has often selected biological concepts arbitrarily. Inheritance is a key area that has always been a focus and it is also a concept that is formally taught in primary school. Given some of the more recent research around inheritance, it seems sensible to assess children's knowledge of this concept using an unbiased and robust methodology that makes use of more in-depth techniques than phenotypic similarity/dissimilarity tasks to form a useful comparison of previous work with which to contribute towards the debate about conceptual development. Likewise, in order to assess progression of inheritance concepts and examine the range of children's understanding on a variety of related concepts, it seems wise to review concepts that are logically related to inheritance and are also part of the NC (DfE, 2014) for primary science. Only then will we be closer in understanding the true nature of conceptual progression in the domain of biology.

Inheritance, biodiversity, and ecology are topics that are currently taught in the NC in KS1 and KS2, which are logically related to inheritance concepts also and yet, have never been investigated thoroughly in the same context. Although evolution is a concept that was not covered by the previous NC (DfE, 2001), the current curriculum (DfE, 2014), includes evolutionary concepts as part of the KS2 primary science syllabus and so its investigation would be timely. These four constructs (inheritance, biodiversity, ecology, and evolution) provide a logical set of ideas to explore the developmental path of concept formation based around inheritance, which has already been fairly well established in the literature. These constructs[1] are explored in more detail in the section below.

Biodiversity:

Biodiversity is a term used to describe the variety of life in the world or within a particular habitat or ecosystem. It is a less well-articulated idea in the research in comparison to inheritance, and little is known about the possible routes of progression in biodiversity concept development. The existence of diversity is typically directly evident to children simply from observing the world around them and thus implies perceptual registration as the start point in conceptual development. This is because early rudimentary concepts about diversity within a particular environment are likely to be based around taxonomy and categorisation. Taxonomy deals with the description, identification and classification of organisms and this is naturally linked to categorisation, which only refers to the latter. An established body of research already implies children have a natural tendency to categorise and so taxonomy is likely to be grasped fairly early on in development (Krombaβ & Harms, 2008; Snaddon, Turner & Foster, 2008).

[1] Note that henceforth 'constructs' refers to only inheritance, biodiversity, ecology, and evolution either collectively or individually.

Of course biodiversity covers a broad range of sub-concepts, which go beyond simple taxonomy and categorisation. However what taxonomy and categorisation allow is the recognition of differences in diversity in the environment and perpetuation of these differences: for example, knowing biological creatures are different from other things and that biological entities give birth to the same kind. By naturally recognising this, it follows that children's early ideas around biodiversity are also likely to be linked to those of biological inheritance, and yet this is something that research thus far has failed to acknowledge on any solid terms.

The basis of acquiring coherent inheritance concepts may result from the fact that inheritance and growth are part of everyday discourse and are observable phenomena. For example, parents frequently comment on the resemblance of their children to themselves (Hyde & Linn, 2006). This, perhaps implicitly, endorses essentialist views about inheritance in the absence of any sophisticated biological knowledge, as Carey (1985) claims. If taken to be true, the same must also stand for biodiversity concepts: grouping animals based on their observable similarities (i.e. taxomy, a key aspect of biodiversity) is something children do very well from a young age as described earlier (Springer, 1999; Keil, 1986; Keil, 1987; Altran & Medin, 2001; Bang, Medin, Altran. 2007) and therefore there is reason to suspect children's knowledge of biodiversity might be better than their knowledge of inheritance, at least initially.

Additionally, Schtulman and Schultz (2008) suggest that knowing a species' identity and categorising it allows one to determine where organisms should live because children's categorisations would have been formed based on observations and clustering related concepts about related organisms. In this sense categorisation is driving knowledge about context, which may then in turn constrain the idea further by essentially forcing children to attend to relevant and associated pieces of acquired knowledge. This is in line with work by Almeida and colleagues (2013).

Conversely, Inagaki and Hatano (2008) claim that children acquire new information about other biological properties of organisms by *anthropomorphising*; first comparing the organism to humans, and then projecting this onto the organism. Carey (1985) stresses the importance of animism being semantically and conceptually different for young children, so much so that they tend to rely on concepts related to psychological properties than biological ones. For example, when asking children "*what is a heart for*" they would often provide a psychological response (e.g. for loving) rather than a biological one (e.g. pumping blood around the body for survival). Interestingly, animistic tendencies such as these do not pose a problem for biodiversity concepts, because it is children's ability to categorise related things together. However, note that with the latter a concept of inheritance is required.

In this regard, not only is biodiversity a potential start point for routes of conceptual progression in biology (because of the early categorisation tendencies of children) but it is also interwoven into other concepts such as inheritance and ecology (Assaraf & Orion, 2007; Hipkins et al, 2008). It is precisely this interconnectedness that past research has failed to engage with and so nothing is yet known about the expected trajectories of developmental progression in these areas.

Ecology:

Ecology is the scientific study of interactions among and between organisms and their physical environments. The idea of *interdependence* is a key feature of the definition. This includes knowledge on the relationships between organisms but *also* the environment; abiotic (non-living factors: water, temperature, food etc) and biotic (living) factors. Ecology is organised under many levels of interaction with individual sub-concepts that need to be linked in order to have a complete understanding. For example, in order to understand the concept of 'ecosystem' (all the organisms that live in a place along with the physical environment), one must first have prior knowledge of 1) species (organisms that can grow and produce fertile offspring) that live in the same area to form 2) a population, and that different populations can live in the same area to form 3) a community which, when interacting with the physical environment, forms an ecosystem (Jordan & Duncan, 2009). Thus logically, one would predict that different levels of ecological concepts are acquired at various ages because of the implied differences in complexity.

A study by Hipkins, Bull, and Joyce (2008) examined children's ideas about ecological phenomena through an interactive drawing task. They found that children would often provide answers to the researcher's questions using psychological reasoning e.g. "the animals will try to all get along (if they live in the same place)" child aged 8, p.74. Thus Carey (1985) argues that although the domains of naïve psychology and naïve biology are separable, a domain of biology arises out of a domain of psychology because prior to holding accurate biological ideas, children have acquired more by way of psychological information. It is only when enough biological information has been acquired and children start to favour biological explanations over psychological ones, that domain of biology is established. In fact findings by Williams and Smith (2010) also support the idea of children's knowledge becoming increasingly biological with age.

Lewis, Driver, Scott and Wood-Robinson (1997) investigated children's (aged 5-16 years) ideas about ecology and found that the majority of children held inconsistent concepts. When asked about various ecological phenomena, children not only held different ideas, but also provided different explanations of the same idea depending on the context they were referring to. Williams and Smith (2010) found similar results relating to inheritance knowledge although arguably Lewis et al (1997) found more fragmentation in children's ecological knowledge. It may be that children often only have 'social' ideas with which to ground their knowledge when a lack of any scientific concepts are available, and as Hipkins and colleagues (2008) say, even though the explanation the child provided is psychological, they have still grasped the basic idea of an ecosystem.

Hence, it may be that with a lack of scientific understanding children rely on the domain of psychology to fill in the gaps and in doing so, promote their understanding of a concept. Thus while Carey (1985) might be right in that children develop a biological domain out of a psychological domain, one would suggest that the domain of biology is simply less scientific, overlapping with a domain of psychology, and later goes on to develop more strongly into a domain of *scientific* biology (Zaitchik et al., 2013). If accurate, the fact that children frequently anthropomorphise would imply that children do not display theoretical understanding, but have naïve ideas which seem fragmented by default (because they are not theoretical but based on overextension), and have concepts that are generally a result of biases.

Similarly, research by Grotza and Basca (2003) indicates that context is highly important when it comes to assessing the theoretical nature of children's ideas. They found that children who demonstrated theoretical understanding of ecological concepts in secondary school tended to have more prior experience with the particular context being assessed than those who did not. Hence the researchers highlighted the importance of familiarity and context in promoting more globalised concepts about causality and concluded that in general, children's ideas are likely to be fairly piecemeal without relevant contextual knowledge. This is further supported by work from Almeida, Vasconcelos and Strecht-Ribeiro (2013) who found that children's past experience of ecosystems were generally in more managed and controlled environments e.g. parks/zoos rather than natural or semi-natural environments e.g. forests. However they note that children's lack of experience of natural environments does not seem to exert any negative influence on their non-anthropocentric reasoning, suggesting any type of experience or exposure is generally helpful for children to develop ecological concepts.

Ideas about ecology need to be fractionated into different elements because a young child is unlikely to have the cognitive capacity to deal with or to understand all of the components of this broad construct as Grotzer & Basca (2003) argue. By separating the large number of ecological sub-elements one is able to assess what kinds of concepts are understood when and what other types of knowledge will influence this. Sander, Jelemenská, and Kattmann (2006) found that secondary school children refered to generally consistent properties which are visible in the everyday environment and that their knowledge of ecological *processes* was far worse. They argue therefore, that the scientific conceptions of ecosystem, imbalance and the dynamics of biodiversity would be difficult for children to understand. For a better understanding, the dimensions of both space and time should be included in curricular design. For instance, ecology encompasses the idea that species eventually adapt to their habitats over time, thus an understanding of inheritance is presumably required to recognise the possibility of successive adaptation over generations to further conceptualise the idea that the environment and organisms are all interconnected in some way to form part of our ecosystem. Of course, this would also feed into basic ideas about evolution.

Evolution:

Evolution is the process by which different kinds of organisms are thought to have developed and diversified. Evolution covers a wide range of complex concepts but Shtulman and Shultz (2008) define two branches: micro-evolutionary concepts including inheritance, variation, and adaptation, and macro-evolutionary concepts including speciation, domestication, and extinction. It is certainly the case that knowledge about evolution would apparently require prerequisite knowledge of inheritance, biodiversity, and ecology to be fully understood. In addition there seems to be a fine line between biodiversity and evolution when notions of adaptation to the environment are considered. A key difference, as has been suggested by many (e.g. Sander et al., 2006) appears to be the temporal dimension to evolution that may make macro-evolutionary concepts, which are usually taught in KS3 and KS4 (DfE, 2011), harder to grasp.

Interestingly, Maurice-Nelville and Montangero (1992) shy away from any ‘evolutionist’ terms and choose instead to refer to the kind of integrative thinking involved: diachronic thinking. They use this term with reference to the ability to construct a series of connected concepts over a temporal axis, which they argue can only occur between the ages of 8-12 due the level of cognitive processing required. Their study examined children's ability to understand forest disease as the disease progressed over time, to eventually be able to reconstruct it, and specifically observe the point at which relatively piecemeal concepts about the disease become integrated into a series of connected ideas about forest disease evolution. This line of investigation is directly linked to the current understanding of evolution in the present research, because the interest is in children's understanding of inheritance turning into a more sophisticated and integrated concept when understood *over time.*

There is virtually no research on the development of evolutionary concepts among primary school children presumably because any work conducted using the essentialist paradigm by default fails to consider aspects of intra-species variation which is a key element to understanding natural selection, the mechanism behind evolution. In this sense, a methodology that captures all the sub-elements in relation to inheritance concepts is likely to help gauge the nature and structure of children's knowledge, which past studies have so far failed to capture.

The majority of the research about children's understanding of evolution has been among secondary school children, and this has largely focused on teleogical reasoning. There has been a wealth of research suggesting children have a *teleological* bias in the way they think about biological phenomena. To have a teleological bias is to believe that category features exist for some purpose and is most commonly found amoung research into children's ideas about evolution. For instance, the idea that giraffes evolved long necks in order to reach food on high trees, in other words the effortful action on behalf of the giraffe is a teleological belief (Keleman, Emmons, Schillac, & Ganea, 2014). Reasoning about categories is informed by how people make sense of the world; this explains why children have nuanced categorisations (Keil, 1999; Hirshfield & Gelman, 1994). For example, Keleman (2004) and Evans (2001) have argued that students learning about Darwinian evolutionary theory frequently display a teleological bias, believing that entities must evolve for a purpose. These findings are even true among expert scientists, though only under speeded conditions (Keleman, Rottmann & Seston, 2013) corroborating the findings of Shtulman and Valcarcel (2012) who demonstrated the retention of naïve theories among adults.

Interestingly, Ojalehto, Waxman, and Medin (2013) suggest that teleological reasoning may instead reveal an understanding or belief of perspective relationships in which nature is an ecological connected web. For example, the relationship between trees and birds could be considered as "*well if you're a bird, trees are for providing homes, food, or shade*" (p.7), hence if there is a purpose for the existence of trees, that must be it. This kind of reasoning may then also account for children's understanding about organism and habitat relationships and other related evolutionary and ecological ideas, thus such a form of reasoning might not be naïve but instead take account of how causal systems function in a more global fashion, thus implying coherent and theoretical beliefs.

Despite young children demonstrating coherent and theoretical ideas about biological phenomena as some would suggest (e.g. Gelman, 2003, 2009; Carey, 1985), it is still the case that children's level of understanding is highly unlikely to reach the same level of understanding to that of an adult's (Piaget, 1970), especially given adults' vastly greater experience. Also, these findings have been obtained from research into secondary school children. As was described earlier, there is reason to believe primary school children may struggle to understand more macro-evolutionary concepts given the temporal axis (Maurice-Nelville & Montangero, 1992; Grotzer & Basca, 2003; Sander et al., 2006; Keleman, Emmons, Schillac, & Ganea, 2014) but may find some of the micro-evolutionary concepts easier. Regardless, the research indicated that children are likely to have limited knowledge and understanding int his area, consequently resulting in teleological biases.

Moving on, a way forward in examining the nature of conceptual progression in the domain of biology would be to investigate a range of related concepts in an in-depth fashion. Findings obtained from earlier work describing theoretical or coherent concepts in young children often made use of constrained paradigms and in some cases, flawed methods. More recent work in this field has begun to use more robust methods and consequently showing children's ideas to be more fragmented than initially thought. The key messages from this recent work suggest that children can hold fairly detailed factual knowledge about biological phenomena, without having much understanding about the causal or mechanistic processes behind them. Hence, children's factual knowledge might be a result of the fact that they are naturally able to make probabilistic judgements from their environment, and successfully categorise information. These ideas are explored below.

2.3 Probabilistic judgements

Probabilistic judgements would also explain results from essential paradigms. A lack of biological knowledge about a particular phenomenon, may force children to rely on teleological, animistic, and essentialist assumptions. The progressive change from social to biological reasoning lends some support to Carey's (1985) thinking that a domain of biology may arise out of a domain of psychology. On the other hand a young child is obviously not taught complex, mechanistic, causal biological explanations about different phenomena and researchers cannot expect to judge children's reasoning as such. A child's knowledge about inheritance at age 5 for example, cannot be thought to be framed in terms of causal biological reasoning because children at that age are not formally taught about genetics, sexual reproduction, or cell biology. For this reason when young children are asked about phenomena of which the causal mechanisms are unknown, they can only offer explanations using knowledge that they have available to them (e.g. Grotzer & Basca, 2003; Sander et al., 2006), and in a case where there is a lack of knowledge, children are susceptible to biases (e.g. Evans, 2001; Keleman, Emmons, Schillac, & Ganea, 2014).

This could be in effect what Carey (1985) describes: children know that species run true from direct observation, though they may not describe it in this way, and as they have a lack of the scientific knowledge and indeed language (scientific or otherwise) available to them, they may describe the process somewhat inaccurately in an effort to capture the weight of probability. However, the exact nature of knowledge and therefore conceptual development is still not captured by this. This is in part due to a lack of research investigating related and in-depth understanding of biological phenomena, but also in part due to a failure to move beyond theuse of highly constrained paradigms that may push children into seeming to say more than they really understand such as the phenotypic similarity task. For example, if children are using a certain means of causal reasoning then they should be using it to explain biological phenomena regardless of the types of questioning.

For example, Strevens (2000) notes that children are not necessarily committed to the view that essences underlie the laws of inheritance. He hypothesises that although children might believe that 'something' about a squirrel causes it to eat bugs, there is not sufficient evidence to suggest that the strength of this inference might be greater if the 'something' was an essence. Hence it could merely be depicting a pattern of statistical inference (Strevens, 2000). In support of this, recent studies have examined the extent to which young children's judgements conform to probabilistic models and found that children are quite capable of detecting patterns, which lead them to infer causal relationships from the age of two.

Schulz and colleagues (2007) looked at children's ability to learn causal structures from the outcome of a series of interventions using a novel toy; an electronic detector with two pegs which would activate when a particular set of coloured gears were placed on the pegs. In one study, children were also shown pictures of a set of gears and that were explained to them e.g. "*The picture shows that yellow is pushing green. Yellow makes green go*." Children were then able to manipulate the gears and at the end of the trials they were asked to select the picture that showed how the toy was working. In the second study children were told how the toy was working and were asked to predict causal outcomes following manipulations. The final study allowed the child to figure out how the toy worked for themselves. The series of studies provided evidence that pre-schoolers aged 24 months were able to accurately identify patterns of evidence from the interventions allowing them to learn the causal structure of events, and in turn allow them to predict the outcome of novel interventions. However the ability to detect patterns of regularity does not necessarily equate to a mechanistic understanding, even of a rudimentary kind, and if children are led into saying something, they may simply be trying to express their sense of observed regularity.

2.4 Categorisation

The tendency for children to make probabilistic judgements also connects with their tendency to categorise this information. The 'Quinnian' (Quine, 1960) theory of categorisation is that learning about natural kinds comes through only a few examples in one's environment in a trial-and-error manner. Quine describes this as the 'problem of induction'. Children expect objects to belong to a kind and generate generic categorisations based on that. Labels are strong identifiers of kind (Csibra & Shamsudheen, 2015) and so it is likely that the increasing use of language would also aid this tendency.

It is apparent that the paradigms used to test children's notions of inheritance often involve unrealistic and over-simplistic scenarios. For example Keil (1996) showed young children pictures of animals and told them about transformations doctors made altering the characteristics of these animals, such as a tiger having its fur bleached and a mane sewn on so that it resembled a lion. Children were then asked if the animal became a lion or was still a tiger and the findings showed that young children maintained that the animal's identity would not change. The reason this seems to succeed as a paradigm might simply be because it abides by the rules of causal laws connecting category membership. Indeed a large body of research has demonstrated that children are naturally good at categorisation from a very early age with some suggesting that even new-borns display primitive categorisation abilities (Mareschal & Quinn, 2001) therefore indicating that children's knowledge of biodiversity may develop the earliest (Krombaβ & Harms, 2008; Snaddon et al., 2008).

Categorisation ability is important because the way children group items together influences the way they learn about the relationships between items and how this knowledge might be generalised. Studies suggest children as young as 3 months are able to make distinctions between domestic cats, tigers, dogs, and birds (Younger & Fearing, 1999). This exhibits how children are naturally making broad distinctions between animal categories and could explain why in the case of biodiversity and inheritance concepts children are able to understand the maintenance of identity (Krombaβ & Harms, 2008).

Gelman (2015) argues that children's knowledge about transformation does not stem from perceptual experiences such as observing regular patterns in the environment; however the experimental examples are sound evidence that children's early knowledge of this kind would be purely perceptual (e.g. Keil, 1994). Likewise, and perhaps more convincingly, switched-at-birth tasks work cross-culturally with animals but not across social categories like gender (Gelman, 2015). These tasks take the idea that off-spring of one or more animals (including humans) are switched at birth and raised by a different type of animal; the child is then asked to decide whether off-spring will grow up to be the same as its biological parents or adoptive parents. Differences in categorisation between adults and children have also been shown in these tasks. For instance, Rhodes and Gelman (2009) investigated whether children shared adults' intuitions that categories vary across cultural contexts and found they did not. Indeed, animals are observable phenomena whilst social categories are arguably not, which would imply children are able to make probabilistic judgements about the former but not the latter. Hence, essentialist tendencies may be indicative of probabilistic judgements.

It may be that the status of research paradigms currently are so constrained, either by using limited methods or exploring limited topic, that they have simply *led* children to provide coherent-like answers and talk about something of which they have no real grasp, other than a natural tendency to categorise. Perhaps children are simply demonstrating their efforts to capture the force of strong patterns of association that they have observed in the environment, and a tendency to tacitly assume, without any particular mechanism being inferred, that these reflect causal laws connecting natural kinds and their observable properties. In reality, it may be that children know very little to begin with and that their concepts are more fragmented as been shown more by recent research (e.g. Williams & Smith, 2010; Williams, 2012). The consequences of this will be explored further in Chapter 4.

2.5 Summary

In sum, much of the research into children's naïve ideas about biology has largely focused on concepts of inheritance. Studies that investigated this in the past argued that children have coherent and in some cases theoretical understanding of inheritance concepts by age 4. However, more recent work is showing another picture, that of fragmented understanding of inheritance. It has been suggested that the reason children seemed to be demonstrating coherent understanding in the past was because of their tendency to observe their natural environment and make probabilistic judgements and categorisations in rather constrained paradigms. In a sense these tendencies are simple heuristics (which may or may not be intentional) that occur as a result of perceptual information. Given that past work only ever looked at children's surface knowledge, the impression was that of coherence, when in reality children have very little understanding because this, as has been found with recent work in the fields of inheritance (e.g. Williams, 2012) and other related areas like ecology (e.g. Sander et al., 2006), biodiversity (Snaddan et al., 2008), and evolution (Shtulman, 2006).

Indeed the early research paradigms used to investigate children's naïve ideas of inheritance are amenable to more simple interpretation: they reflect probabilistic judgements and categorisation. Therefore, in order to fully comprehend how children are learning naïve ideas about inheritance, we ought to use a more robust and unbiased paradigm that uses these two variables, and not essentialism, as the point of departure.

Similarly, in order to understand conceptual change in naïve biology, other related concepts to inheritance such as biodiversity, ecology, and evolution should also be tested in order to see how knowledge in one area might influence knowledge in another, in order to observe the points at which isolated biological knowledge might be incorporated and integrated. This becomes particularly important if it is the case that children's understanding is fragmented because establishing a sequence of conceptual development would enhance teaching practices and curricula considerably.

CHAPTER 3

3.1 Conceptual Change

Research in the fields of education and psychology has shown that children enter school with existing ideas about how the world works and that these prior conceptions are resistant to change in certain contexts (particularly those relating to more socially prominent phenomena to a greater extent) despite not being fully developed or accurate (Driver et al., 1985; Howe, 1998). Studies have also demonstrated that the misconceptions children have are often very resistant to change, even when children are provided with explicitly contradictory, although accurate, information (Howe, Tolmie & Sofroniou, 1999; Fugelsang & Thompson, 2003). This phenomenon poses a problem for effective instruction; it becomes important to understand how children are acquiring and organising knowledge and how this ultimately leads to conceptual change and progression.

Kuhn (1962) introduced the notion of conceptual change to indicate how concepts are embedded within a theory and can change alongside any alteration of the theory itself. Specifically, Kuhn (1962) argues that the process of concept change occurs as a result of an accumulation of anomalies and a resulting crisis, which essentially drives forward a change in the conceptual structure. Although this is a philosophical perspective, it could still relate to science learning in children by comparison. Indeed, conceptual change has been recognised as a fundamental aspect of science learning (Mayer, 2002; diSessa, 2006) because learning science depends on more than just a simple accumulation of facts, but rather on a child's ability to restructure existing ideas on the basis of any new content and experience of the everyday world (Zaitchik, Iqbal & Carey, 2013; Vosniadou, 2002). This poses a problem in the context of teaching because learning science is likely to involve conceptual change much of the time, and pedagogical strategies need to take this as a general point of departure.

As discussed in Chapter 2, the process of conceptual progression has only been explored in limited fashion, with the focus tending to be around the way children organise the information they acquire. There are many different models of organisation, which makes theorising about conceptual change difficult. In cases where children are able to successfully acquire accurate biological concepts, it is assumed conceptual change has occurred. However, there is still debate about exactly what the process of change entails both in the short-term and long-term. The difficultly is compounded by the fact that the nature of conceptual *progression* is unknown and under-researched.

Theorists in this area generally discuss conceptual change within two broad perspectives: those that view knowledge acquisition of novel material only in the context of formal instruction, and those that consider the reshaping or restructuring of existing knowledge via experimental mechanisms.

Even within these two camps of conceptual change, there have been largely two schools of thought that dominate the discussion about how conceptual change comes about. Some have argued that knowledge is represented in a theory-like structure and highly organised, with concept change occurring through theory-replacement processes (e.g. Vosniadou, 2002, 2014; Vosniadou & Brewer, 1992, Gelman, 2009; Carey 1985). Others argue concepts are not theoretical, but a collection of fragmented ideas that are context-specific, with development more iterative and concepts more malleable (e.g. diSessa, 1993; Harris, 2002).

The differences of opinion imply different pathways towards how biological knowledge is structured, and therefore the consequent educational implications are likely to vary. For example, there is a lot of literature regarding conflict and change, reconstruction of knowledge structures through peer collaboration and inhibition rather than reconciliation (discussed in more detail below). However, the difficulty with all of these explanations is that they all seem to have evidence in support of them, suggesting different processes might occur in varying contexts and conditions. These theories of conceptual organisation and resulting change are now discussed.

3.2 Accounts of conceptual change

3.2.1 Equilibration

Piaget (1971, 1972, 1985) paid much attention to the development of children's concepts of physical and biological phenomena. His theory of cognitive development claimed that children were born with an innate set of mechanisms that allowed them to represent, store, and use information they gained from their environmental experiences. By this account, Piaget views conceptual change as reshaping existing knowledge structures, which children obtain by acquiring perceptual information from their environment. Children are first able to *assimilate* this information, and try to incorporate this knowledge into existing schemata or modify the schema entirely by a process of *accommodation*; by doing so, children reach a state of *equilibration* between internal representation and external reality and between related internal representations. However the process of knowledge acquisition and conceptual change is reliant on the child encountering conflict, that is to say if a child encounters a novel scenario which their current model of conceptual knowledge cannot explain, a state of *disequilibration* is reached resulting in a series of transformations occurring to the schema, so that it becomes progressively adequate.

In Piaget's explanation of conceptual change, knowledge results from continuous construction of new structures that did not exist before, either in the external world or in the child's mind. The development of scientific thought, according to Piaget's framework, is a process of continual organisation and reorganisation. Sophisticated concepts emerge via processes of coordination (e.g. the same action performed on different objects leads to coordination of the constant aspects into a more general action concept) and differentiation (e.g. different types of outcome to the same action will result in more refined action-outcome concepts). This type of approach also implies that the underlying mechanisms by which children learn different portions of information are essentially the same i.e. the processes and mechanisms to acquire inheritance concepts would be the same as those required to acquire ecology concepts.

Piaget (1970) stresses that there are many different schemas about perceptual material encountered in ones' environment that can be coordinated with one another, thus implying a general system of coordination of perceptual information. It is the resulting schemas that are the point of departure for logical structures. Theories that push this framework of learning are known as domain-general theories, where there is an assumption that development leads to a structural change (Gelman, 2009). The process of sophisticated and on-going refinement of knowledge structures, as described by Piaget (1970), is necessary because no single logical schema is sufficiently strong to support all knowledge that is encountered. This may explain why even within area of naïve biology such as inheritance, children can have well-developed knowledge about phenotypic similarity with quite poor knowledge about phenotypic dissimilarity despite the fact that they are heavily related concepts (Williams & Smith, 2010).

The consequence of this is that when all the different logical schemas a child has are taken together, they are not sufficiently coherent with one another to serve as the "foundation for human knowledge" (Piaget, 1970, p.10). This would suggest that children's ideas about biological phenomena are not theoretical, particularly at the beginning of knowledge acquisition in contrast to early inheritance research (e.g. Springer, 1999) and therefore the degree of integration among different but related biological concepts would not be sufficiently high to yield a theoretical and coherent understanding. This issue is also highlighted in Piaget's work where children's understanding about conservation of area was acquired sooner than conservation of volume because, according to Piaget, it involves fewer dimensions, and conservation of number sooner still for the same reason (1970). Research exploring Piaget's notion of conservation (Prince-Williams, Douglas, Gordon, William, Ramirez & Manuel, 1969) found that some children did acquire conservation of volume earlier because of direct experience suggesting that experience rather than complexity is key. Prince-Williams et al (1969) recruited 28 6-9 year-old boys from pottery-making families in Mexico with 28 matched controls and hypothesized that boys who were experienced in pottery-making would demonstrate earlier conservation skills (particularly around the concepts of weight, liquid, and volume) than those boys who had no pottery-making experience. Results confirmed Piaget's hypothesis, ultimately suggesting that direct experience of phenomena is more important than the simplicity or complexity of that phenomena. If this were the case, then examining the nature of input would be key in trying to understand the nature of rudimentary knowledge and how it is stored.

Similarly, as Hipkins et al (2008) and Almeida et al (2013) note, it may also mean examining the contexts in which children have acquired knowledge as it would indicate children's initial conceptions are likely to be heavily context-specific.

Vygotsky (1962, 1978) was one such researcher who examined the nature of input a child received. He believed language plays a critical role in children's conceptual development. Social interactions drive cognitive processes including language, thought, and reasoning because the linguistic social interactions (external speech) the child has with adults later become internalised as thought (inner speech). This approach suggests that language is a key driver for conceptual change, given the co-construction of knowledge that takes place.

However, this approach does not clarify with any specificity how children develop inaccurate intuitive biological theories, and seems to consider conceptual change only in the context of formal instruction, unlike Piaget. By default, if a child is internalising what an adult has explicitly demonstrated through speech, they should develop fairly explicit and accurate ideas from a relatively young age. In this regard, Piaget was certainly pioneering in suggesting that conceptual development is not solely the product of formal education, but involves something of the internal construction of knowledge, and although he strictly dealt with the *explicit* content of children's knowledge, he was still able to provide some valuable insights about what the origins of implicit naïve biological concepts might be that eventually lead to explicit theoretical concepts.

Nevertheless, this model does not directly highlight the ways in which children coordinate and integrate their knowledge of different biological constructs. One model that potentially accounts for this is proposed by Karmiloff-Smith (1992), described below.

3.2.2 Representational Redescriptions

Following on from Piaget's work, other perspectives on conceptual growth suggest that ontological experience (basic categories of being and their relations) is one area of input that is key in terms of coordinating fragments of knowledge that may be implicit, into more explicit and broad structures (e.g. Karmiloff-Smith, 1992; Carey, 1985; Inagaki & Hatano, 2004, 2008). Karmiloff-Smith's (1992) model of *representational redescription* (RR) asserts that conceptual development begins with the acquisition of fragmentary and implicit representations of action-event relationships that are triggered when appropriate circumstances are encountered, but are otherwise not mentally malleable. She argues that the human mind has both detailed innate specifications and predispositions, as well as skeletal domain-specific predispositions that are context-dependant. In other words, she implies that environmental input is necessary for development and in the case of detailed innate specifications; the environment simply acts as a trigger for the organisation of knowledge, whereas in the case of the skeletal outlines, the predispositions are not detailed and in these instances the environment is likely to act as more than just a trigger, but as a key influence for subsequent conceptual organisation and potential change.

This is an interesting idea and has similarities to a model of conceptual change offered by West & Pines (1984; 1986) who attempt to explain the full gamut of knowledge conditions by considering the reshaping and restructuring of existing knowledge prior to or after instruction. They argue that in any instance where children have no prior intuitive concept, and are formally taught new material, children must attempt to acquire symbolic knowledge from their environment in order to make sense of their reality. In this situation, which the authors note is typical of school biology lessons, learners are often forced to ignore their own reality whilst trying to integrate symbolic knowledge. This, they argue, is the process of conceptual change.

West and Pines (1984) highlight that the differences in understanding between student and teacher knowledge might hinder or enhance learning; children who are taught material that is consistent with their existing ideas are able to learn effectively, whilst those who are taught material that is inconsistent with their existing ideas struggle to understand new concepts and are often resistant to the suggestion that their own ideas might be inaccurate. A fault line with this model however, is that conceptual change is only viewed in the context of formal instruction, and the exact mechanisms which might take place allowing conceptual change have not been specified. As a remedy to this, Karmiloff-Smith (1992) claims that knowledge acquisition is brought about when the mind exploits the information it has already stored internally (both innate and acquired), by iteratively redescribing its representations of this information or "*re-representing in different representational formats what its internal representations represent"* (p.15). This entire process aids the child to be able to turn implicit ideas into explicit concepts, allowing conscious construction of knowledge and theoretical manipulations of it. Therefore this model indicates that not only is knowledge likely to be context-specific as others (Hipkins et al., 2008; Almeida et al., 2013) have also stated, but that children have very fragmented initial concepts which they acquire from direct observation and gradually build up and refine over time to develop more coherent and integrated biological ideas.

The RR model postulates different representational formats at different levels: I (implicit), E1 (explicit 1), E2 (explicit 2), and E3 (explicit 3). These are parts of a reiterating cycle that occur within different domains throughout the developmental span. At the Implicit level, information is embedded in procedural form, new representations are independently stored, and the potential links between domains remain implicit. The processes of representational redescription for phases E1-E3 compress the procedurally acquired formats in the implicit level. At the E1 level, explicitly defined representations can be related to other redescribed representations, unlike at the implicit level. Representations at E2 have conscious access but are not yet available to verbally report, hence they are partially explicit and E2 is mainly a stage for extra representational redescriptions of those at E1 with increasing levels of explicitness. Only representations at level E3 are verbally explicated and these representations are accessible across domains and communicable. This view is in line with Piaget (1971) who argued that acquisition of schemas gave rise to linguistic structures (one possible form of symbolic representation). The RR process also provides a useful explanation as to why children have different levels of understanding in related biological areas (e.g. Williams, 2012) as Piaget (1971) does, but goes further in explaining why even older children and adults still hold naïve biological theories (Williams & Smith, 2010; Shtulman & Valcacel, 2012).

Karmiloff-Smith (1992) claims that attentional biases and some innate predispositions make one susceptible to focus on linguistically relevant input and with time, build up domain-specific linguistic representations that eventually become modularised. She argues Piaget's theory of sensorimotor development (Piaget, 1971) does not explain how language acquisition is initiated and argues that in the case of language, the mind becomes modularised as development proceeds (Karmiloff-Smith, 1992). Her account also considers brain plasticity in early development and she suggests environmental input is constrained by initial biases or predispositions that channel attention to relevant environmental inputs, which in turn affect subsequent brain development. This suggests a potential route for how newly acquired language might influence subsequent conceptual change constraints. Hence Karmiloff-Smith in her account provides further detail as to how general cognitive abilities may also contribute towards children's domain-specific knowledge acquisition. This is discussed further in Chapter 4.

The premise that human cognition involves domain-specific mechanisms is a common idea. Other cognitive developmental psychologists also view domain-specific learning as a result of the constraints on learning (e.g. Keil, 1987). Domain-general and domain-specific theorists treat the idea of constraints on development differently. Constraints for the former would act as a factor to curtail a child's competence, for example with Piaget's theoretical stance, constraints would limit the overall system with which to acquire, assimilate, and accommodate environmental information, which would subsequently curtail successful conceptual change. This would ultimately mean that conceptual progression of children's biological knowledge would be exceptionally inefficient and piecemeal, as mentioned earlier. In contrast, constraints enhance learning in domain-specific models such as Karmiloff-Smith's, by limiting the number of extraneous variables and focusing only on relevant pieces of information with which to strengthen children's representations from implicit to explicit levels. For example in the RR model, verbally explicable concepts at the E3 level are only formed as a result of the redescribed perceptual representations acquired at the implicit level. Hence the input representations constrain the output representations. This highlights the ways in which children are likely to have domain- and context-specific knowledge when it comes to biology (Assaraf & Orian, 2003; Grotzer & Basca, 2003) and hence, also why children have varying degrees of understanding among related biological phenomena (Krombaβ & Harms, 2008).

Keil (1996) describes constraints as restrictions on the kinds of knowledge structures that the learner typically uses in order to restrict and guide understanding, and these constraints could either be acquired developmentally, or be innate (Keil & Lockhart, 1999). In the case of the RR model (Karmiloff-Smith, 1992), constraints are innate and subsequent environmental input constrains output. This is an interesting idea because it may be that a child is learning something within a particular context and thus the context itself may be the constraining factor (e.g. Hipkins et al., 2008). Assuming domain-generality, children should be able to learn certain scientific concepts at the same rate as other related concepts and be equally proficient in both. However, this does not seem to be the case. For instance

Tolmie et al (2009) have shown that even within a single curricular topic such as *physical state change*, understanding of melting among 8-year-olds is not predictive of understanding of evaporation. Likewise in the field of biology, work into children's understanding of animate versus inanimate knowledge would logically rely on the same grasp of the distinguishing properties of animate entities and inanimate objects, yet work into natural kinds suggests otherwise (Gelman & Markman, 1986). This would imply the type of contexts children are exposed to can restrict or *constrain* the theoretical formulations leading to domain-specific conceptual change.

The process of RR that Karmiloff-Smith (1992) describes is the same across domains, however this does not imply that there is simultaneous change across domains, just as Piaget (1970) also claimed. The difference with the RR model is that it is a process by which implicit information in the mind becomes explicit information of the mind, first within a domain, and then subsequently across domains. In the first level of the RR model, children focus mainly on environmental input hence this stage is data-driven, much like Piaget (1970) would argue.

The RR model attempts to account for the ways in which children's representations become more flexible and malleable, to eventually allow the emergence of conscious and explicit knowledge for theory building. This allows the child to form *representational adjunctions* that neither alter existing stable representations, nor are brought into relation with them. Thus indicating that concepts develop laterally where children may use knowledge of one biological area to help them understand another. Successful performance can be generated by a sequence of independently stored representations that will ultimately have to be linked into a more coherent system. Hence level one is the only stage to have environmental input where the subsequent changes and redescribed representations are internally driven. At level two, the temporary discard for environmental evidence may lead to errors and inflexibilities. In phase three the internal and external data are reconciled, for instance in the case of language, new mapping is made between the input and output representations in order to restore correct usage (Karmiloff-Smith, 1992). Once new representations are stable, they are added, domain-specifically, to the existing concepts with minimal effect on what is already stored. Therefore independently stored representations do not cultivate representational change. This suggests that knowledge in this phase is very piecemeal and is merely an accumulation of various domain-specific environmental inputs, at least initially, but over time, with explication and coordination of biological concepts, knowledge may indeed be theoretical.

This view is in line with Gelman's (2009) account that conceptual development is the product of *both* innately specified mechanisms and also through environmental experience. This argument is persuasive because domain-general theories, such as Piaget's (1970) assume there is no element of mechanistic cognition during learning; the focus is on learning and learning output, rather than the cognition that takes place in between to allow learning to occur. The RR (Karmiloff-Smith, 1992) model postulates that domain-general processes sustaining inference and representational redescription operate throughout development (and are likely to be innately specified) and in order for representations to go across different domains, a process of explication that is aided by increasing linguistic ability must undergo representational redescription. In contrast to the Piagetian framework, Karmiloff-Smith's RR model will spontaneously improve itself *even in a state of stability* upon environmental exposure and provides a better explanation of conceptual change over time than Piaget's account where conflict and *disequlibration* must be encountered in order for change to occur.

The exposure to different ontological experiences (relationships between different categories) serving as a guide to constrain children's conceptual change is something also claimed by Carey (1985). Carey argues constraints in the environment allow children to acquire relevant information so that learning may in turn be sequential. When children's innate predispositions are not sufficiently specified whereas Karmiloff-Smith (1992) argues the environment acts as a key influencer of change, Carey (1985) maintains children are able to pursue intuitive knowledge across domains, in a theoretical manner, until more knowledge is acquired.

3.2.3 Carey's account of conceptual change

Carey (1985) is one of the few researchers that have developed an account of conceptual change that specifically relates to children's biological knowledge. Carey (1985) argues that domain-specific knowledge emerges from different ontological experiences a child is exposed to, which go on to restructure children's concepts rather than endorsing the view that domains require distinct ways of acquiring knowledge through innate mechanisms. Unlike Piaget, Carey (1985) proposes that conceptual change cannot be considered as global restructuring, but domain-specific restructurings of naïve theory structures on the basis of increasing new knowledge about a particular biological concept. Hence, as children are exposed to new experiences and instruction, they gradually replace their naïve theory-like structures with scientifically correct conceptual structures.

Similarly, Chi, Slotta and DeLeeuw (1994) offer an explanation about how conceptual change might occur in the context of acquiring new biological knowledge through instruction. They claim that some concepts would be harder to learn than others and posit a model of conceptual development named the *Incompatibility Hypothesis*. They propose that in a classroom setting causal relationships cannot be understood until the categories with which to organise new knowledge are well defined, suggesting that it is the organisation of new information with prior related information that is the basis for understanding. Chi et al (1994) view conceptual change as a concept redefined from one category to another. In other words, conceptual change is a category shift, which they argue causes a change in the ontology of a particular concept. By default, this model is able to account for why conceptual change for some biological concepts might be easier to organise and understand than others, because if a child is introduced to a concept that cannot be readily organised into a pre-existing group of related ideas, difficulty arises. However, unlike Carey this model does not fully consider the actual processes involved in conceptual change or conceptual progression, nor address the fact that children arrive in school with pre-existing ideas that may influence restructuring or transformation of ideas rather than simply be used as a path map to conceptual organisation.

According to Carey (1985) restructuring occurs through processes of replacement (one fundamentally different concept replaces another), differentiation (the initial concept splits into two or more new concepts e.g. the concept of cat can be split into Persian, or tabby), and coalescence (where two or more concepts coalescing into a single concept e.g. Persian and tabby into one concept of cat).

Carey (1985) claims that children's early beliefs about biological phenomena are couched in psychological domains, which can often lead to misconceptions about biological kinds. For example, Carey (1989) quotes one child as saying "worms don't have babies, they just have little worms" as demonstrating that this child views "having babies" in terms of the social roles of parenting (psychological domain), and not in the reproductive sense (biological domain). In this way knowledge acquisition might even be seen as sequential in nature whereby explanations for biological phenomena become more 'biological' as more biological knowledge is acquired. A similar view is also taken by Hatano (1990) where prior to any formal instruction, children are in possession of well-developed biological ideas that allow them to make consistent predictions and form explanations regarding biological phenomena. At first a child's ideas are uncoordinated but then eventually develop within the domain they belong to, increasing in complexity. This, Carey (1985) says "*typifies all novice-expert shifts regardless of age*" (p.71). Hence, one way children address a problem is to assess the degree of similarity a phenomenon has to an exemplar they already have knowledge about, thereby allowing domain-specific theoretical changes.

Recent studies offering support for the domain-specificity of children's theories (e.g. theory of mind; theory of physics) make apparent that content knowledge is integral to structure and development (Gelman, 2009; Wellman & Gelman, 1998; Carey, 1985). Carey (1985) claims that children organise their learned concepts into coordinated systems and beliefs such as physics, psychology, and biology, as shown by the degree of variation in sophistication of ideas sufficient to drive explanation. These ideas probably derived from the differences in ontological experience children have had, and the exposure to different ideas.

Tolmie (2012) notes that both explicit explanation and implicit beliefs have been regarded as an index of children's understanding, and yet the actual origins of these explanations or beliefs and their relationship with conceptual structure have been unexplored, as has evidence suggesting that children's learning is heavily context-dependent and subject to familiarity effects. For example, Inagaki (1997) claims that familiarity with a domain of knowledge increases not only the amount of factual knowledge children have but also their conceptual organisation of that domain. Hence, Carey (1985) argues that a lack of familiarity with concepts in the domain of biology results in children using existing concepts in the domain of psychology to make sense of their world.

The organisation of knowledge has so far been the main focus with regards to how children are learning from the environment, for example, Rhodes and Gelman (2009) propose that children may believe that both animal and artefact categories reflect true structure because they associate with groups that share similar properties. In this sense categories are "*natural bundles*" in the environment (p.247) i.e. they are ontologically different as Carey (1985) would suggest. This provides further support that children might be context-bound with regards to scientific phenomena, as category boundaries are often applied in domain-specific ways.

Chi and colleagues (1989) are also of the view that domains can be seen as guides with which to categorise things for example, although there are a number of ways to classify living things, some beliefs about living kinds are typically early emerging and consistent. Thus domain competency facilitates this by focusing attention on a specific domain rather than general knowledge (Chi et al., 1989), which both Carey (1985) and Karmiloff-Smith (1992) argue children are predisposed to do.

Indeed, feature extraction and association is what children do extremely well (Gopnik, Sobel, Schulz, & Glymour, 2001), and categories/schemas emerge naturally from that process, as discussed in Chapter 2. It has been argued that causal information can serve to organise other information in a way that makes it easier to recall (Grotza & Basca, 2003). For instance when memorising facts about birds, an expert will be able to grasp the causal (i.e. cause-and-effect) relations between facts (e.g. birds can fly because they have wings). Causal reasoning might be largely domain-general, its properties are largely independent of topic area but conceptual understanding is seen as domain-specific with marked topic variation of levels of sophistication and the degree to which it is sufficiently explicit to drive an explanation. Studies have shown that fostering the use of causal relations helps novices to understand biological facts and to recall the information over a longer period of time (Opfer, Nemh & Ha, 2012). This view is similar to that of Carey (1985) who argues that a biological domain can only ever be attained if deductive reasoning (i.e. using multiple premises to logically guide reasoning) is attained, presumably by understanding the causal relations between phenomena, however she also mentions the fact that adults do not consistently use deductive reasoning, which would imply that causal reasoning might not necessarily be explicit. This suggests that harnessing causal and deductive reasoning might be a domain-general strategy that could be generalised across a variety of different topics.

What seems to be evident is that there is a natural tendency for children to extract causal information from experience because of its organisational benefits, and this may explain why children are sensitive to co-variation and causation from a young age (discussed further in Chapter 4), possibly because of an innate predisposition to do so. Piaget (1970) also emphasised children's structural knowledge is experientially driven, as does Karmiloff-Smith (1992). If the external environment is chief in the initial organisation of domain structures, these domain structures themselves must be a constraining factor in the development and change in children's concepts. Hirschfeld and Gelman (1994) note that if knowledge was left unconstrained, instruction would rarely yield meaningful knowledge, unless by chance. For instance, within a topic such as evolution there are constraint-based interactions crucial to a complete understanding (Chi et al., 1994), suggesting that it may be easier to acquire simpler related concepts, which would subsequently constrain the child. This would also imply that a complex biological concept such as evolution might dip into other ontological categories in order to be fully understood, potentially organised in some kind of network of related ideas within a domain (Vosniadou, 2014).

In fact Carey (1985) argues a domain of biology arises out of a domain of psychology for this specific reason: children have limited biological knowledge early on and in an effort to understand certain biological phenomena, they will often resort to psychological explanations. For example Myant & Williams (2005) investigated children's understanding of health and demonstrated that in the absence of knowledge about the mechanisms involved in illness, children would refer to behavioural or psychological explanations, but that explanations would become more biological with increasing age. This ultimately implies that conceptual change in children is theoretical in nature. Others (Gelman, 2003; Gopnik & Wellman, 1994) also have similar views such as Carey (1985) who believes a domain of biology arises from a domain of psychology through restructuring theories about biological phenomena. Indeed Gopnik and Wellman (1994) claim children's early conceptions in the domain of psychology are implicit theories and any changes in those conceptions are theory changes; children are like "*little scientists*" (Gopnik & Wellman, 1994, p. 259).

However, while it is certainly true that limited biological knowledge will lead to limited conceptual understanding in this domain, the fact that children go *across* domains to explain unknown biological phenomena would suggest that theoretical domain-specific conceptual integration would take longer to achieve and perhaps be more susceptible to error, given the lack of a foundation of rudimentary biological concepts.

Vosniadou (2002, 2014) puts forward an account to explain how children might coordinate pieces of related information in a theoretical manner based on different ontological experiences that is similar to Carey's (1985).

3.2.4 Vosniadou's framework account

Vosniadou (2002) argues that children's concepts are embedded in a theory-like structure and subsequent formation and change is theoretical. Vosniadou and Ioannides (1998) argue that children's beliefs are linked to and consequently constrained by ontological presuppositions, which allow concepts to form within a coherent structure. These are presuppositions about the kinds of entities children assume to exist and the way that they are organised. For instance, one assumes that in their ontologies there are entities such as physical objects, which can either be categorised as animate or inanimate. Of those that are animate, children can further assume that animate entities have certain properties themselves (Vosniadou & Ioannides, 1998). In this regard, the only difference between children's naïve beliefs and adult's expert beliefs, are the ontological presuppositions children may make about explanatory mechanisms, which become refined over time.

Thus this framework theory approach is able to account for children's difficulties in learning counter-intuitive science concepts, such as the Earth is spherical and not flat as it appears, for example (Vosniadou, 2014). However, exactly how concepts relate to each other has not reached any consensus. It is precisely for this reason that the role of a domain in conceptual change and development might be especially salient with regards to constraining, integrating, and coordinating existing and subsequent knowledge within a child's framework.

Vosniadou (2014) maintains that domain-specific learning occurs through formulating hypotheses about the ways in which specific content is structured and restructured, without the need for innate constraints or modules, but instead via the use of domain-general mechanisms. There have generally been two major foci in this regard; the first is cognitive conflict, and the second is meta-representation (Vosniadou & Ioannides, 1998). The former is the idea that children have pre-existing concepts, which either corroborates with any new knowledge they learn or does not, i.e. the two concepts (old and new) are in conflict with each other (e.g. Piaget, 1970; Chi et al., 1994; West & Pines, 1986). The child must then resolve this conflict either by choosing the existing or novel concept, or form a new concept altogether, hence cognitive conflict causes cognitive adjustment.

However little has been said by Vosniadou to explain the processes involved in conflict resolution in the context of this specific framework, unlike Piaget (1970) who offered the account of equilibration described earlier. Consequently, without being able to understand the mechanisms behind concept change in biology one will never be able to understand the process or developmental trajectory of conceptual progression, or indeed what, how, and why category shifts occur. Hence this framework theory in the context of cognitive conflict appears to be under-specified.

The other process or cause of conceptual change is the development of meta-representations. This is the notion that children will hold multiple ideas or concepts in their mind in one large network of related concepts, yet within this network would exist the initial naïve concept as well as its more sophisticated form (Vosniadou, 2014). For instance children may hold the naïve idea that inherited phenotypic traits are a result from sharing a home with their parents, to the more sophisticated biological idea that sharing parental genetic material leads to similar phenotypes. Similarly, Gelman (2009) argues that biological concepts are not isolated fragments rather they are part of an integrated network of larger, related concepts, culture, and other idiosyncrasies much like the ideas of Vosniadou (2014) and Carey (1985).

However, this does not address the cause of resistance to new ideas in children; one would assume that if children were capable of holding multiple representations of concepts in their mind then they would also find including new information into their web of knowledge relatively easy. Conversely, it could be that meta-representations can only be formed when related ideas cohere together to form a theoretical framework (e.g. of biology), and if a concept does not relate to existing concepts in child's network, it is rejected i.e. ontological presuppositions determine the new information that a child acquires in order to develop a meta-representation of existing material. In the RR model (Karmiloff-Smith, 1992), innate predispositions constrain children's representations through a process of representational descriptions. Though again, Vosniadou fails to adequately explain how concepts might be coordinated within theoretical frameworks.

Shtulman and Valcarcel (2012) found evidence supporting framework theories as described by Vosniadou (2014), and go as far as to say that these framework theories may remain until adulthood. They found evidence of the maintenance of naïve theories of natural selection, even among biology experts, when under speeded conditions. This suggests that there may be inferential competition between theories rather than the replacement of one theory with another (e.g. Tenenbaum, Kemp, Griffiths, & Goodman, 2011), as the knowledge-as-theoretical supporters (e.g. Vosniadou, 2014, Carey, 1985) would argue. This view would also explain how cognitive conflict resulting from a mismatch in prior beliefs of children to new material might be resolved.

In their study, Shtulman and Valcarcel (2012) gave college participants a speeded task on a number of topic areas, including natural selection knowledge, where participants responded true/false to statements. They found that participants were more accurate at responding to 'consistent' statements (where naïve and non-naïve statements were both true/false) rather than 'inconsistent statements' (where statements were true for either naïve or non-naïve theories and false on the other). They also found a difference in response times between inconsistent and consistent statements was significantly correlated with accuracy scores, and interpreted these results by suggesting that students with greater domain-general expertise might have exhibited more cognitive conflict on the inconsistent items than did participants with less expertise. Results also suggested that naïve theories are suppressed by non-naïve scientific concepts but are not supplanted by them, possibly because naïve theories emerge earlier and are therefore more deeply entrenched (Shtulman & Valcarcel, 2012). Their results coincide with the meta-representation view that Vosniadou (2014) describes whereby conflict does not *cause* conceptual development, rather the ability to hold multiple concepts in an organised framework does.

Alternatively it could be that under limited processing speed and limited resources, students resort to using a heuristic of sorts. Perhaps naive theories and scientific theories occupy different levels or have different forms of representation (Shtulman & Valcarcel, 2012). The former may take a more situational and context-specific representation, whereas the latter may be more abstract and general. This is similar to the RR model where Karmiloff-Smith (1992) argues concepts develop over different levels to eventually become explicated and more sophisticated. If so, then the perspectives offered by Shtulman and Valcarcel (2012), and Vosniadou (2002, 2014) regarding theoretical conceptual development would actually point to theoretical integration of initially acquired fragmented concepts. This view echoes sentiments of Harris (1994) who believes children undergo processes of hypothesis testing and have working models that serve as a basis for prediction and explanation. However, Harris (1994) stipulates that learning in this sense is not theoretical, but dependent upon concrete paradigms, where the child is thinking like a working scientist piecing together fragments of information.

3.2.5 diSessa's account of conceptual change

Some have argued that a common-sense or folk theory, based on intuition rather than fact, is not the same as a scientific theory (Karmiloff-Smith & Inhelder, 1974; Carey, 1985; Keil, 1994; Gopnik & Wellman, 1994) because everyday thought may not be theory-like in its resistance to counter-evidence, ontological commitments, attention to domain-specific causal principles, and coherence of beliefs (Hirschfeld & Gelman, 1994). Indeed it seems as though the lack of consensus has led to mixed opinion about the definition of theoretical thinking, and if it were the case that a folk theory is not the same as a scientific theory, this would imply conceptual change is not as theoretical as previously thought, and may be far more fragmented. By default, this would also suggest that children's beliefs at an initial stage are unlikely to be coherent either, not least until they are able to be sufficiently coordinated or verbally communicable as Karmiloff-Smith (1992) describes.

The main champion for the 'knowledge-as-pieces' view is diSessa (1983, 1988, 1993, 2006). diSessa's viewpoint is that knowledge structures consist of various conceptual elements at different stages of development and sophistication, which have been constructed bottom-up out of sensory and perceptual primitives, or p-prims. From this perspective concept change is a piecemeal evolutionary process, rather than a theory replacement process, where p-prims lead to the development of mental models that will iteratively change given new context-specific knowledge. These p-prims are loosely organised in a conceptual network, and specific p-prims are activated by recognition mechanisms under exposure to relevant contexts. Hence, knowledge is highly context-specific. However, p-prims are universal in their function as diSessa describes them, which seems somewhat contradictory to the idea that they should result in fragmented knowledge.

diSessa (2006) suggests that conceptual change results from a cognitive reorganisation of the naïve knowledge elements into more complex systems, thus with the acquisition of more elements an eventual theoretical framework is possible. For instance, fragments of ideas on plant biology could be loosely grouped together with fragments around ecology, which after appropriate activation in relevant contexts, could lead to a more holistic understanding about ecosystems (e.g. Krombaβ & Harms, 2008). With this view, diSessa's (1988) theory would also help to explain exactly how children go from implicit to explicit knowledge via appropriate triggers in the environment. This view is very similar to that of Karmiloff-Smith (1992) with regards to pieces of knowledge becoming more coordinated as they become more explicit. Relevant contexts, and the increasing coordination of linguistic units allows explication of rudimentary and implicit concepts to occur, which eventually might lead to a more theory-like conceptual view once a certain level of coordination has been reached. This would create a level of conviction that Piaget (1971) refers to as necessary knowledge, whereby logical reasoning is necessary in children's thinking in order to identify a species of knowledge. In this sense, necessary knowledge is a pervasive feature of every domain where necessary conclusions allow children to go beyond the information they obtain from the environment (Smith, Laughran, Berry, & Dimitrakopoulos, 2012). Hence it seems as though there are certain similarities between the two major perspectives.

3.3 Knowledge as fragmented and emergently theoretical

Ozdemir and Clark (2007) conducted a study aiming to find support for one of the two perspectives (knowledge-as-theoretical or fragmented), and found support for the knowledge-as-pieces perspective overall after replicating an earlier study of Ioannides and Vosniadou (2002). In the study, the meaning of force and its development among 105 children was investigated. It was found that 88.6% of children's responses fell into seven categories of internally consistent interpretations of force: internal force, internal force affected by movement, internal and acquired force, acquired force, acquired force and force of push/pull, force of push/pull, and gravitational and other forces. The students in the study gave consistent predictions and explanations regardless of context, which the authors concluded as confirmation that children have uniform and internally consistent interpretations of force.

However, diSessa et al (2004) conducted a direct replication study of Ioannides and Vosniadou's (2002) and found children's meaning of force were inconsistent across contexts. Most importantly, diSessa et al (2004) found evidence that children's knowledge about force was heavily context-dependent, where pieces of knowledge were activated under certain contexts (diSessa, 1993). Ozdemir and Clarke (2009) conducted a replication of diSessa et al's (2004) study with Turkish children, and found overall support for the knowledge-as-pieces view. They highlighted that methodological flaws such as soft coding schemes and limited contexts could lead to overestimations about student's knowledge and the coherence of student's knowledge.

Likewise, Southerland, Abrams, Cummings and Anzelmo (2001) conducted a quasi-replication of diSessa et al's (2004) study and examined the structure of children's biological knowledge. They found that a knowledge-as-pieces view had more explanatory power than a knowledge-as-theory view, because p-prims could better explain the shifting nature of students' conceptions of biological phenomena over the frameworks theory (Ozdemir & Clarke, 2007).

This work is in line with more recent research suggesting children may have fairly fragmented ideas (e.g. Williams, 2012; Williams & Smith, 2012) and many are now of the opinion that the mechanisms allowing a child to learn are domain-general, yet the topic-specific variation in their knowledge would imply domain-specificity in content. The idea that knowledge is domain-specific suggests that many cognitive abilities are specialised to handle specific types of information (Hirschfeld & Gelman, 1994). As such, it is highly likely that children acquire fragmented topic-specific concepts, and it is not yet clear how these concepts develop, or are coordinated to lead to more holistic understanding about particular phenomena. The RR model proposed by Karmiloff-Smith (1992) makes an attempt at describing how fragmented concepts become coordinated over time via a process of redescribing implicit representations to explicit ones that are verbally communicable. If taken to be accurate, in the mental models diSessa describes there is still the possibility to hold a theory-like model which will iteratively change, where the role of language is central to the explication and coordination of ideas over time.

In fact, this process of explication and coordination of fragmented concepts goes as far as explaining the retention and inhibition of naïve concepts. Concepts that are not matched with a linguistic unit will not reach the final E3 stage of explication (Karmiloff-Smith, 1992), likewise any concept that has not been through the necessary iterations of redescribing representations means that concepts might remain implicit, or only partially explicit resulting in the maintenance of naïve theories. In any instance where representations are redescribed on the basis of new external information, previously held ideas are inhibited in favour of more coherent and robust concepts.

Indeed the evidence for the retention and inhibition of naïve concepts is vast and any model attempting to credibly explain conceptual change should be able to successfully account for this. Given this point, the knowledge-as-pieces view seems to be the most sustainable argument. Vosniadou (2014) maintains that initially children acquire fragments of knowledge but these form a relatively coherent explanatory system, or framework theory, which allows for prediction and explanation of new phenomena. The idea of a framework theory almost places Vosniadou in middle-ground territory as on the one hand she accepts thoughts can be relatively coherent, and yet the framework is not necessarily explicit or has any internal consistency as a scientific theory would. It is also not a socially shared theory, which implies that domain-specific knowledge may be highly individualistic and dependent upon on the early experiences and observations of the child. What a framework theory is able to do according to Vosniadou (2014), is be used to predict and formulate hypotheses and generate explanations. Hence domains like naïve biology and physics may be less closely linked with our scientific understanding in the formal sense of the word than is sometimes appreciated (Vosniadou & Ioannides, 1998).

However, Vosniadou (2014) claims that framework theories are not explicit, and that conceptual change is a gradual process. If taken to be true then, framework theories cannot be theoretical in the scientific sense as this would require them to be explicit, as Karmiloff-Smith (1992) argues, echoing the sentiments of diSessa (1993). Hence, on the surface, the two views are not completely incompatible, however the key difference is that Vosniadou assumes coherence (at least locally) and this cannot account for why children perpetuate and maintain naïve theories as shown by Shtulman and Valcarcel (2012) and also other studies such as that by Howe, Tavares, and Devine (2012) who demonstrated that children's responses to tacit and explicit trajectory tasks are completely at variance. The accounts offered by diSessa (1993) and Karmiloff-Smith (1992) are better able to explain results such as these because explication and coordination seem to be relatively piecemeal. The implication of this then is that fragmentation can be quite extreme, especially at the perceptual or explanatory boundaries, and coordination only happens through relatively deliberate and conscious effort.

3.4 Interim summary

Overall it would seem that even though there is some consensus that children's knowledge and conceptual change is theoretical (Carey, 1985; Vosniadou & Brewer, 1992; Vosniadou, 2014; Gopnik, 1996; Gelman, 2009), much of the accounts and perspectives offered by the relevant theorists seem not to be sustainable as a position. The domain-general account by Piaget (1970), which has influenced many conceptual change models, is unable to explain exactly how children coordinate concepts theoretically. Moreover, the very fact that children have separate logical schemas for different phenomena means that children must have different rates of understanding for seemingly tightly related concepts (e.g. children's understanding about conservation of area was acquired sooner than conservation of volume). Similarly Karmiloff-Smith (1992) offers an account of how conceptual change might lead to eventual theoretical understanding, but not before acquiring separate fragments of knowledge. Even in a specific domain, or a specific area within a domain there could be markedly different apparent levels of progression, for example Tolmie and colleagues (2009) showed children's understanding of melting is plainly ahead of their understanding of evaporation and condensation throughout the primary age groups and others have shown similar findings in the domain of biology (Gelman & Markman, 1986; Williams, 2012). This suggests that the level of development is likely to vary for different domains of knowledge, and one cannot assume that cognitive development will be the same across multiple or even related domains. Hence even if ostensibly tightly connected concepts have uncertain relationships to each other, then the nature of the connection between taxonomically related concepts certainly cannot be assumed, but needs to be empirically demonstrated.

It seems as though the fragmented knowledge accounts of conceptual change seem to offer a more viable position than the theoretical knowledge conceptual change accounts, both in the way of explaining how change from naïve to more sophisticated concepts might take place, but also in terms of explaining the evidence from studies about the retention and maintenance of naïve theories. Both views stipulate that the acquisition of concepts is highly context dependent, and this is supported by other research (Assaraf & Orian, 2003; Hipkins et al., 2008; Almeida et al., 2013). If so, then instruction should focus on how these fragments of knowledge are activated in appropriate contexts, perhaps by developing an incremental curriculum, which would allow for the organisation, modification, coordination, integration, and refinement of biological concepts over time.

As discussed in Chapter 1, many researchers have not been overtly clear about what they mean regarding concept development or change. Domain-general theorists accept the notion that all types of concepts change in the same way, but very little consensus has been reached about the nature of conceptual progression and it is not yet clear whether conceptual change involves restructuring existing knowledge, and development involves encompassing new concepts *into* existing knowledge as Piaget (1985) and West and Pines (1984) would suggest. The many different models of organisation ultimately make theorising about conceptual change a difficult task.

Both perspectives of the knowledge-as-theory and knowledge-as-pieces debate would agree that children observe phenomena in their environment, which in turn influences their conceptual knowledge, but there are differences as to how sophisticated this knowledge is. Some have suggested that the way initial concepts are organised goes on to constrain future acquisition and organisation of new concepts. Consequently, investigating the strategies and mechanisms behind knowledge acquisition should also be examined (Siegler, 2000). However, an issue with this is that a failure to address the nature of change means that there is not adequate evidence to inform theories about the mechanisms involved. The account offered by Karmiloff-Smith (1992) seems to be the most convincing in that is it the development of language that allows the coordination of atomistic ideas into larger structures. It could be, however, that with the gradual coordination that language offers, concepts may become more theoretical over time. In any case, initial conceptual change certainly seems less likely to be theoretical.

3.5 Language

Gelman (2009) takes the view that language helps children to coordinate their ideas, and serves to organise their concepts in such a way that a combination of explicit testimony, implicit cues on language, and child expectations and capacities work together to guide conceptual learning. Gelman (2015) however, argues that the type of language children are exposed to matters a great deal with regards to how children are able to categorise information, and has shown that children are sensitive to the use of generic language or lack of. For example Brandone and Gelman (2013) found generic language use reveals domain differences in children's expectations about animals and artefact categories. In their study, 5-year-old children and adults' use of category-referring generic noun phrases (e.g. "birds fly") were examined for novel animals and artefacts. Results revealed children and adults produced more generic language when items were described as animals, as opposed to artefacts, even though the stimuli were perceptually identical. The authors attributed the findings to participants' ontological expectations about animals and artefact categories. Indeed, particularly in the context of science within a classroom setting, specific language will refer to specific phenomena for example "a bird is an animal that can fly" yet this generic phrase does not consider flightless birds which, according to Brandone and Gelman

(2013) children would categorise as flying animals based on their sensitivity to generic language. Gelman (2015) argues children begin to use generic language about natural kinds at around two years of age despite limited observations, hence this level of abstraction would suggest theoretical learning, as Carey (1985) would also claim.

One issue to consider here is the difference between language use in context, and language use in testimony that is removed from context. Both are important but testimony requires a degree of prior elaboration of existing concepts in order to be effective. Therefore, language use in context may be more important to begin with, helping to direct attention toward important features, making connections between these and providing sign functions that enable these to be explicitly manipulated. For example, Philips and Tolmie (2007) demonstrated that providing 8-year-olds with explicit explanations about balance scale problems promoted learning only when the children had a foundation of basic knowledge about the role of weight and distance i.e. externally input language must match the child's existing level of functioning to some extent. Without this basic knowledge, learning more complex material created confusion. This is similar to the view that linguistic units have to be redescribed to the same level as conceptual units in order to be coupled together explicitly in the RR model (Karmiloff-Smith, 1992). Hence it seems trying to promote higher-order reasoning skills before initial precursors are established may hinder learning.

One idea about the influence that language has on conceptual development is that the labelling used in everyday language helps associate the word to a perceptual feature and aids the categorisation process that children are so good at (Csibra & Shamsudheen, 2015). These lexicalisation effects are automatic consequences of the fact that the label serves as an automatic cue to influence a child's judgements. However there are situations where the child may hear a word yet fails to learn it, thus lexicalisation effects are not automatic as the child must also make an evaluation and judgement. This would also highlight exactly how within the right context, children’s fragmented concepts could be activated as diSessa (1993) posits in the knowledge-as-pieces view. Indeed, the influence of language brings up important questions to consider. Is it that language serves as a mechanism with which to bring together ideas and create new concepts (Gelman, 2009), or is it that language can draw attention to and highlight other available concepts and thus promotes conceptual progression?

Fulkerson and Waxman (2007) suggest that the process of learning words is essentially matching (and understanding) a linguistic unit to a conceptual unit. Early studies investigating inheritance concepts of young children often failed to consider the importance of language in conceptual knowledge, as the methods used in those early studies often required children to make simple judgements (e.g. Keil, 1997; Springer, 1999; Gelman, 2015). The reason for these early language-sparse methods may have been because much of the research was conducted on pre-schoolers who arguably are unable to articulate themselves. However, more recent research into this area suggest that children as young as 4 are able to demonstrate quite sophisticated ideas about inheritance and other biological phenomena using speech (Myant & Williams, 2005; Williams & Smith, 2012), and even though children are still required to make judgements in some of these tasks, they were asked to explain the judgements they made in an interview-type task, which aided researchers in understanding the limitations of their knowledge.

However, one thing to consider is the extent to which different approaches might be reconciled to produce same or similar results. In the previous chapter it was argued that language-sparse approaches may have led to children appearing as though they had more advanced conceptual knowledge than if they had taken part in a language-heavy approach but this does not necessarity mean that children should only be given language-heavy tasks as language-sparse approaches may be informative of children's implicit conceptual understanding. However, the RR model postulated by Karmiloff-Smith (1992) would suggest that explicit concepts are formed *out of* implicit strands and not independently of them. Hence there ought to be a degree of overlapping or carry-over, even if explicit concepts have become more organised. This however, does not mean that language-sparse and language-heavy tasks are addressing the exact same parts of the concepts children hold, but it does suggest value in encouraging children to give explicit responses sooner rather than later because the building blocks already exist. Therefore, it appears sensible to use tasks that require both judgement and explanation and any future research in naïve biology should consider the importance of including a language-based element to their method, even when examining younger children.

Evidence suggests that infants as young as three months old are able to form language categorisations to non-human primates' vocalisations (Ferry, Hespos & Waxman, 2013). It is only when infants begin to understand word meaning that they are able to map that learned word to a specific category across varying contexts. It may then be suggested that learning the meaning of novel words in different contexts could take longer to understand than when the same context is used consistently. For these reasons, priming would have to be sufficiently powerful to deal with the unstructured nature of everyday events, and thus cognitive conflict with peer collaboration may help increase the salience with which primes are held in memory. The authors suggest that learning from collective insights (a Vygotskian view) and learning from subsequent events may be two sides of the same coin; both are dependent on opposing ideas and with the balance resting on the degree of contradiction and the child's skill to deal with this (Howe, McWilliam, & Cross, 2005). This hypothesis explains how peer collaboration induces cognitive conflict to promote concept change *over time*, specifically via explicit language used to provide causal explanations of particular phenomena.

The work into peer collaboration also highlights the importance of language as a potential supporter towards conceptual development. In fact parents might also be socialising their child to think in more relational ways (Gelman, 2015), hence increased social relations might play a role in the construction of knowledge. This again, is a very Vygotskian idea but may also explain the benefits of collaborative group work observed in other studies (Howe, Tolmie & Rogers., 1992). Ganea and colleagues (2007) further suggests that one function of collaboration through language is to allow a child to update knowledge and beliefs in the absence of any direct experience with the object or phenomenon itself, which also explains why peer collaboration might promote concept change.

3.5.1 Peer collaboration

It is often assumed learning is a solitary act when in actuality it is usually embedded within a social context (Gelman, 2009). For example studies of theory of mind suggest certain types of learning require attending to others as a key source of information (Baldwin, 2000 as cited in Gelman 2009). This of course is not a new idea, and based around earlier Vygotskian theory. Vygotsky (1978, 1983) held the view that the mind is a network of specific and independent capabilities, which allow for the ability to think about a variety of things. However as Siegler (2000) notes, despite this early belief, the initial research into children's development was more focused on learning rather than the thought processes behind that learning. Studies into collaborative learning have suggested that this may be a specific context in which language use appears to play a key role in engendering productive conceptual shifts, by allowing the coordination of fragmented ideas as Karmiloff-Smith (1992) suggests.

Williams and Tolmie (2000) studied 8- to 12-year-olds' ideas of inheritance in a peer collaboration setting. Children were pre-assessed individually and placed into peer groups with either similar or dissimilar ideas to their own, and took part in a task designed to promote conceptual conflict and group discussion. Post-test interviews with individual children showed greatest conceptual advance among those children who were placed in groups of peers who had dissimilar ideas to their own, suggesting that encountering high levels of conceptual conflict promoted dialogue in these groups, leading to conceptual change. Similar findings of collaborative group-work studies have also been shown in the domain of physics, although often the conceptual benefits appear after some time.

A study by Howe and colleagues (2005) examined why the beneficial effects of peer collaboration are not always apparent until time has elapsed by investigating children aged 9-12 years on their understanding of floating and sinking objects using a peer collaboration task. They found that delayed effects were not down to any post-collaborative reflective appraisal, nor the breakdown of any unhelpful representations over time, rather that peer collaboration primed children to make better use of subsequent related observations and evidence. In their study, Howe et al (2005) provided demonstrations highlighting the different factors towards objects either floating or sinking, which were viewed as probes, and demonstrated fortnightly. They gave these to one group of children and not the other, and then gave all participants a collaborative group task. The results suggested strong evidence that peer collaboration can prime children to make productive use of subsequent events and observations in a way that “chance seems to favour the prepared mind”. The study suggests the mechanisms include primed sensitivity to useful events and that priming might depend on cognitive conflict because if a child who had incorrect prior knowledge, collaborated with someone with differing views, and was therefore exposed to new vocabulary, experiences, or explanations, this would prime them to assess and think about any future encounters that they may have. In this way, it could be that cognitive conflict and the resulting priming from collaborative group tasks promotes conceptual change within individuals. Note that this might not actually be a necessary element, just one that might be helpful because it results in stronger traces.

3.6 Accounts of mechanisms involved in collaborative learning

Studies researching the impact of collaborative learning, point toward the central role of transitive dialogue that appears to be consistent across a range of studies. Howe et al (2007) conducted a study investigating children’s understanding of condensation, evaporation, floating, and sinking in group-work projects where children had to propose ideas, explain their reasoning and resolve any differences of opinion with other members of the group. Their study found that knowledge gains were predicted by the proposition/explanation variable in group-work contexts. The frequency for disagreements was also higher in group-work contexts, although disagreements themselves were not consistently associated with knowledge gains. The authors suggest that it may be disagreements create a context where propositions and explanations are more likely, which serve as indirect support to learning.

Based on Piaget's account (1971) of conceptual change, it would seem as though the reason collaborative learning is successful is because children are exposed to new information that comes into conflict with their existing knowledge. This causes a state of *disequilibration* by which children are forced to assimilate and accommodate existing schemas. For example in the case of Howe and colleagues (2005) who conducted a study similar to Williams and Tolmie (2000), children were asked about their predictions for materials that would float or sink and the reasons for why. Children were asked to make predictions about what objects they believed would float or sink, prior to confirming whether or not all children within a group agreed. There were two conditions: those where children had similar ideas and so agreement was easy, and those where they had differing ideas and so agreement would have been more difficult. Children then participated in a practical experiment to identity which objects floated or sank, and were asked to discuss the results in their groups until everyone in the group could agree about what caused objects to float or sink. Those children whose predictions did not come true would not be in a state of cognitive conflict according to Piaget. In the study, children within the group had to agree on a potential theory of floating/sinking, which essentially forced the child to accommodate new information (the ideas of the other children that seemed accurate) to eventually form a new schema, thus allowing conceptual change. Thus, for collaboration to drive conceptual change, children must have differing ideas to begin with, in order for cognitive conflict to occur. These findings are of course similar to those of Williams and Tolmie (2000).

Studies of shared learning suggest that reconstitution of knowledge, i.e. re-shaping or organising knowledge on the basis on new information to become more scientific, might also be a difficulty in learning science, and to this end collaborative learning may be beneficial (Howe et al, 2007). This remains to be fully tested but resonates well with the knowledge-in-pieces account of representational development. Indeed the co-construction of knowledge is by no means a new idea, Vygotsky (1978) considered the reconstitution of knowledge to occur in a shared state, where children learn first at a social level, and then on an individual level given the results of collective insight. Under group work conditions, children have different sensitivities to different phenomena and have to marry that with language. It is this mapping with language that provides a basis for manipulation of information needed for scientific thinking and argumentation. Hence Vygotsky (1978) would claim the reason peer collaboration works is because children are communicating ideas with each other which forces them to use and coordinate linguistic structures that eventually allows them to form a coherent argument, and thus a coherent concept.

This explanation by Vygotsky (1978) echoes the sentiments of Karmiloff-Smith (1992) somewhat. The mapping of language onto implicit concepts allows these implicit (or partially explicit) concepts to be borne out allowing coordination and integration of concepts, promoting change. Taking the same study (Howe et al., 2005), The RR model (Karmiloff-Smith, 1992) would indicate that children initially have partially explicit ideas (perhaps at E1 or more likely at E2 level) about the reasons for floating and sinking, as they are able to make some sort of prediction, albeit a wrong one. When the results of the practical experiment go against their predictions, children engage with dialogue with their peers, which as Vygotsky (1978) also claims, allows mapping of ideas onto language. This allows partially explicit ideas at E2 level to eventually reach E3 level where ideas can be coordinated and verbally communicable, thus engendering conceptual change. A problem with this, however, is that the movement from E2 to E3 levels is unlikely to occur right away, but is more likely to occur over time.

A study by Howe and colleagues (1992) investigated the delayed effect of peer collaboration and demonstrated that a delayed effect was not simply down to extra time for consultation with teachers or adults, but rather because of individual *cognitive activity* (Howe et al., 2005). Children are essentially *cognitively* consolidating information alone and it may be the case that the activity of group work facilitates this. This study provides some support of the RR model, therefore.

However, children cannot rely on peer collaboration alone, particularly in instances where it seems as though their peers have inaccurate concepts. Tolmie (2012) considers this issue and highlights the importance of language as a mechanism for conceptual change, but emphasises the conjunction of dialogue with manipulation and observation. Tolmie's two-systems hypothesis suggests children are predisposed to detect all perceptual information, forming something of a tacit knowledge system, similar to Karmiloff-Smith (1992). This system over time and with the aid of language acquisition to coordinate and explicate the tacit ideas into more explicit ones creates a conjunction with the perceptual system and linguistic structures. Tolmie's two-systems account, although convincing, has not been investigated whereas the previous accounts of Piaget (1970; Vygotsky (1978), and Karmiloff-Smith (1992) have studies in support of each of their ideas as presented earlier in this chapter. This would indicate that each account is capturing at least some of what is occurring by way of conceptual change, and yet neither account is fully able to disprove another. The account by Tolmie (2012) attempts to integrate the former three accounts of Piaget, Vygotsky, and Karmiloff-Smith and although it has yet to be tested, the two-systems account seems to resonate well with the idea that knowledge seems likely to be initially fragmented, and that these fragments of knowledge are coordinated by the increasing mapping of linguistic structures to tacit perceptual units, leading to conceptual change into explicit (and perhaps eventually theoretical) knowledge structures.

3.7 Summary

While there is increasing agreement that learning is likely to be domain-specific with underlying domain-general mechanisms, the debate as to whether knowledge is theoretical or fragmented has no clear consensus. The beginning of the chapter compared these accounts and outlined what the implication for the nature of science learning might be. Studies have provided support for both perspectives, however arguments for the notion that conceptual change, at least initially, is fragmented seemed to be more sustainable as a position, given the increasing evidence for this view as described in Chapter 2, also. What was clear from these studies is language seems to be a potential route to aid the process of conceptual change, particularly with regards to differing contexts. One avenue that research in science learning has suggested this might occur is though collaborative learning. Howe et al (2005) argue that little is known about the exact benefits of peer collaboration, which studies have shown can often be delayed but yet results from such studies can be interpreted by Piaget's account of cognitive conflict (Piaget, 1970), Vygotsky's account of the co-construction of knowledge (Vygotsky, 1978), and Karmiloff-Smith's model of representational redescription (Karmiloff-Smith, 1992). The fact that each of these accounts is able to account for the relative importance of language use in collaborative learning suggests that each is only capturing a limited part of the full picture of conceptual change. Tolmie (2012) presents an account that attempts to integrate the three models by suggesting the two-systems hypothesis, which although untested, corroborates the evidence that knowledge is initially atomistic and conceptual change relies on the ability to coordinate fragmented concepts with linguistic structures to reach explicit knowledge status.

CHAPTER 4

4.1 Non-essentialist theories

Chapter 3 illustrated that the nature of conceptual change and progression in naïve biology is still unclear. The majority of research has been embedded in the essentialist paradigm, which as the previous chapter demonstrated, is flawed. The Chapter 2 highlighted that research findings from such studies may be better explained by other non-essentialist theories, as will be described here.

There has been a wealth of research suggesting that children detect and learn from patterns of probabilistic information (Kushnir, Xu & Wellman, 2010; Kirkham, Slemmer & Johnson, 2002). There have been studies (Koerber, Sodian, Thoermer, & Nett, 2005) that have shown preschool children can understand the relation between co-variation and causal belief, but only when one causal factor co-varied with an outcome. Yet when more than one variable co-varied with the outcome, children failed to interpret patterns of empirical evidence (Kuhn, Amsel, & O'Loughin, 1988). It is possible that this pattern of findings relates to the development of domain-general cognitive mechanisms, which might not be sufficiently well developed to coordinate multiple concepts. However, it does not necessarily mean that children may fail to detect patterns of regularity *implicitly*. Explicit theories are thoughts and representations that are verbally accessible whereas implicit theories are those that are not, and yet might still be covertly influential as they are often supported by representations (Zaitchick et al., 2013).

The idea that children detect perceptual patterns from the environment relates to the account diSessa (1993) offers about conceptual change. In this account diSessa claims the perceptual or sensory primitives, or p-prims, that children acquire means that knowledge is piecemeal until context-specific mental models are developed over time. Karmiloff-Smith (1992) takes this account further and suggests that the way in which implicit pieces of perceptual information (or p-prims) are coordinated is by a process of representational redescription which allows implicit ideas to become explicit concepts when linguistic units have been mapped onto them.

4.1.1 Implicit & explicit thought

A study by Howe and colleagues (2012) investigating children's understanding of object fall showed the differences in implicit recognition of object fall and explicit prediction on the same task. They argue that prediction of the object fall trajectory requires children to actively and explicitly engage with their conceptual knowledge, whereas simply viewing different falling trajectories and selecting the one that looked the most accurate (i.e. as in the recognition task) is merely implicit, with no necessity to engage with explicit knowledge. Children in three cohorts across primary age range (Years, 2, 4, and 6) took part in this study. Children were shown a series of computerised scenes involving a hot air balloon in two tasks: recognition, and prediction. In the former, children were told a ball would be dropped from the hot air balloon (that was either stationary or moving) and watched computer-presented scenarios before being told to judge which of the scenes depicted an accurate motion. Three weeks later, the same children were told a ball would be dropped from the hot air balloon again (stationary or moving), and were asked to predict the trajectory in the prediction task by selecting the appropriate route from a series of three potential routes (note that the order was counterbalanced). Results demonstrated that the prediction task was considerably more challenging than the recognition task for all cohorts. In particular, the stationary scenarios had a 61.3% success rate on the prediction task whereas recognition task performance was close to ceiling. With the moving scenarios, success rates were 2.6% on the prediction task and 55.2% on the recognition task. It may be that the gap between prediction and recognition results from omission at the explicit level of elements that are tacitly appreciated (Spelke, 1994), which Howe and colleagues (2012) argue could account for a plausible model of conceptual development.

Interestingly, the moving task was harder for children, despite the fact that the scene may have been more in line with their everyday observations. It may be that there was simply too much visual stimulation that caused confusion. However, the pattern of responses was quite different for the tasks too, with backwards trajectories (always the incorrect answer) being predicted increasingly more often with age, but no such change was evident in the recognition task responses.

Rather surprisingly, it was found that in the prediction task, children in Year 6 performed consistently worse than children in Year 2 where performance was generally more accurate, despite the fact that in the recognition task, the majority of children from all cohorts performed satisfactorily. This suggests that although prediction requires explicit engagement with conceptual knowledge, recognition is achievable through tacit processing. The findings from the prediction task would also suggest that children in Year 2 are relying on their tacit knowledge, whereas those in Year 6 seem to be dismissing their tacit beliefs in favour of a learned behavioural response. More fundamentally, this research suggests that with age children also have a dichotomy between implicit and explicit systems of knowledge (Howe et al, 2012).

Providing support from the field of neuroscience, Kallai and Reiner (2010) conducted an EEG/ERP study that examined adult males' electrical activity immediately following a stimulus event. Participants were shown video simulations of the classic ball trajectory problem by McCloskey and Kohl (1983) on separate trials where the ball would either exit a spiral tube or a straight tube. Participants were then asked to say whether the trajectory of the ball exiting either tube (which would either be a straight or a curved route upon exit) looked accurate or not. In this sense the paradigm was very similar to the recognition task used by Howe et al (2012), who would argue participants would not have to actively engage with conceptual knowledge. Findings revealed that the majority of participants tended to select the curved trajectory as the correct route when the ball was exiting the spiral tube.

However, there was an activation peak at N400 (associated in general with a perceived semantic violation) for the curved trajectory, regardless of tube type. These results indicated that the curved path, which is always the inaccurate trajectory and yet the preference for the majority of participants, always violated expectations at some level given the peak in electrical activity at N400. However this ostensibly implicit violation of expectations was always overridden by an opposing behavioural response in the case of the spiral tube. This strongly suggests that implicit and explicit systems operate separately to some extent, and that in this instance, the implicit system seems to be more accurate.

Interestingly the pattern of responding in this study, which is essentially tacit, is very similar to the pattern of responding found in the prediction task in Howe and colleagues' (2012) study, which is essentially explicit. This would imply that the more explicitly reasoned responses seen in children eventually become relatively automated. Indeed there is some evidence of inhibition of inaccurate conceptions by adults (Shtulman & Valcacel, 2012). The findings from both Kallai and Reiner (2010) and Howe et al (2012) would also imply that there are three layers of representation: accurate implicit conceptions, inaccurate partially explicit conceptions, and accurate fully explicit conceptions. There is not any indication that these three types of conceptions become combined, but rather that they are maintained separately and that the level of representation might be specific to certain contexts. In which case, this is evidence strongly in favour of the knowledge-in-pieces view advocated by diSessa (1993).

Studies using blicket detectors, a device that flashes or makes a noise (manipulated by the experimenter) when a particular configuration of objects, or blickets, are placed on top of it (e.g. Gopnik et al., 2001) also suggest young children are very good at conceptually inferring patterns from their environment, implying that children are relatively good at picking up low-level perceptual information. Studies using this paradigm generally involve manipulation of more than one 'causal factor' at a time and have shown that children's responses show more or less exact statistical sensitivity to the witnessed covariation, unlike other instances where if more than one variable co-varied with the outcome, children failed to interpret patterns of empirical evidence (Kuhn et al., 1988). It might be that once children have established these patterns implicitly, they begin to manipulate the variables involved in their environment, yet as this involves the perceptual system, there is still no formal conceptual understanding.

Consequently if this low-level perceptual system is accurate then one must consider where explicit thought comes from. Perhaps it derives from social contexts (Harris & Koenig, 2006; Rhodes & Wellman, 2013), which would lend support to Vygotskian (1978) ideas of learning where dialogue is the key mechanism behind the distinct overlay of explicit concepts onto previously implicit ones. Indeed the two-systems hypothesis proposed by Tolmie (2012) also echoes these ideas. This hypothesis suggests that an anticipatory system exists in the brain that detects all perceptual information, almost like a tacit knowledge system, which over time and support through tasks that create a conjunction between the perceptual system and language, eventually overlays implicit idea with explicit concepts. What needs to be considered, however, is whether this pattern differs according to different contexts, and if it does, there is likely to be a dual-system model of learning, which might suggest the ways in which children's concepts change developmentally.

It may be that learning is largely implicit where some pieces of knowledge can be formed explicitly straight away. This implies that the prior building blocks of that knowledge, such as domain-general foundations, had to have existed previously for this to work. If this account is taken as accurate, it would seem as though adults and children alike would have accurate low-level systems which cumulatively extract perceptual regularities in the environment to serve anticipatory mechanisms. Considering the RR model (Karmiloff-Smith, 1992), it may be that only some aspects of (implicit) perceptual experience become more explicit through the support of language, at the E1 to E3 stage in the RR model, and become more coordinated as a result of this, leading to actual concepts. Where this happens in a bottom-up fashion, the result is slowly emerging with relatively accurate partially formulated ideas (e.g. the effect of slope angle on motion). Therefore implicit ideas can form a number of explicit (relatively related) concepts, and reach a further level of sophistication when explicit knowledge can reach a level of abstraction, much like Piaget (1985) might have thought.

However, although this is essentially a bottom-up process, it is also highly likely to have top-down feeds for instance, the influence of media, conversations, and teaching. It is likely that occasionally top-down feeds may lead to quite distorted conceptions about everyday phenomena and these distorted ideas may be further reinforced by a priming effect (see Howe, 2006) and eventually automated (e.g. as seen in the backward trajectories of falling objects experiments by Howe et al., 2012 and Kallai and Reiner, 2010). In this account, language seems to be the likely driver for explicit concepts, but it operates in an uncertain fashion and will not necessarily be helpful in the development of accurate conceptions. For this to occur, language would need to be systematically introduced alongside observation to be sure of being productive, which is something that could potentially be introduced under formal instruction at primary school.

Consistent with the layering account is an fMRI study conducted by Fugelsang and Dunbar (2005) investigating physics concepts. They discovered that the anterior cingulate cortex, an area associated with cognitive conflict, was activated when participants viewed a scientifically accurate movie clip that conflicted with their beliefs, as does the dorsolateral pre-frontal cortex, an area associated with executive function. This implies that the task in the study forces observation into merging with prior concepts, where otherwise the observation would have just been glossed and depending on the relative strength of this connection, concepts may be partially accurate, or fully accurate. The response conflict was taken to indicate that students still had access to naïve theories despite conceptual change already having taken place, a finding also observed for evolutionary concepts (Shtulman & Valcarcel, 2012). Perhaps then, it is possible that different mechanisms might be involved in learning, the neurological basis of which is unknown. It would also lend support to Tolmie's (2012) two-systems hypothesis in that explicit concepts overlay implicit ones, implying that implicit beliefs may never truly be lost.

A more recent fMRI study (Mason & Just, 2015) assessed the neural mechanisms behind incremental learning of college students in physics knowledge, specifically mechanical systems. The study was able to show the sequence of neurally-identifiable knowledge states, which the authors argue could be generalised toward other areas of science, including neuroscience and biology. The study suggested that early learning is very piecemeal, with students' focus on specific components to the mechanical system, but with further incremental training, these components could be grouped together in order to focus on the system as a whole. This is in line with Tolmie's (2012) argument that repeated support through tasks helps to create a link between the perceptual system and language, eventually overlaying implicit ideas with explicit ones. Mason and Just (2015) argue therefore, that learning is the transition from one state of piecemeal knowledge to another that encompasses all prior learned components, thus providing further evidence that is consistent with the idea that complete integration of concepts is unlikely to occur, rather partial selection and integration is more plausible. This suggests that incremental and repetitive learning seems to aid accurate knowledge acquisition, and if the same were shown in younger children, this would have inevitable curricular implications. What this study by Mason and Just (2015) does not do however, is demonstrate exactly how children go from having implicit to explicit beliefs, or how the influence of social factors, such as language, are likely to influence children's learning.

As noted earlier, there may be three layers of representations: accurate implicit conceptions, inaccurate partially explicit conceptions, and accurate fully explicit conceptions. These three types of representation are maintained separately and that the level of representation might be specific to certain contexts in that the selection about what pieces of knowledge become explicit and coordinated is likely to be a process driven by specific contexts. Also, the kinds of activity and the timing of exposure support the active and predictive use of context, suggesting that by understanding how contexts provide constraints on ambiguity we might be able to understand its influence on conceptual development. Indeed Skipper (2015) argues that brain regions supporting language are not fixed, and as the guided use of language may aid explication of some implicit ideas, it follows that brain organisation might shift around as a result of contextual information and prior experience, in which case one is likely to expect a developmental change with regards to the kind of contextual information children are using in comparison to the kind that adults are using.

So far the perspectives of Karmiloff-Smith (1992), diSessa (1993), and Tolmie (2012) have led to the account that language may have a role to play with regards to overlaying implicit scientific concepts with explicit ones. Given this account, implicit concepts form as a result of acquiring perceptual information and eventually coordinating this information, which points to a relationship between science learning and domain-general capabilities. As discussed earlier, contextual differences within aspects of science learning would imply that children may have different levels of representations (accurate implicit, inaccurate partially explicit, accurate fully explicit) even among areas of related biological ideas. Perhaps more importantly however, is that if the process of explication is influenced by factors such as media, formal instruction, or everyday conversations about biological phenomena, it becomes important to investigate the domain-general capabilities that coordinate this information. One would assume that better domain-general capabilities would lead to better coordination of ideas, resulting in a higher likelihood of accurate biological conceptions.

As discussed in Chapter 3, if the goal is to assess how children are linking together related biological concepts to form more sophisticated concepts over time, one cannot simply investigate conceptual progression and growth of these concepts alone, but also seek to investigate the domain-general mechanism/capabilities which contribute toward the coordination of piecemeal biological concepts in the first place. Hence one must assess how far any sources of shared variance between different (biological) conceptual areas are simply explicable in terms of general capabilities because any variance that is not explained by these domain-general capabilities, may imply conceptual integration/progression of some kind, which would be more informative for research given the current nature of curricular sequencing.

4.2 Importance and influence of general cognitive functions

The development of children's knowledge is likely to be dependent upon the increasing sophistication of domain-general capabilities. This intrinsically forms the foundation of learning or cognition, but also highlights the importance of investigating domain-general capabilities in children. Indeed the extent to which there are points of connection between different biological areas could be *driven* in part by these underlying capabilities, as discussed earlier.

4.2.1 Numeracy and Literacy

There is already a wealth of evidence to suggest that general cognitive functions influence other areas of learning such as literacy and numeracy, for example Cragg and Gilmore (2014) note that although it is clear that domain-specific numerical skills and knowledge are important factors for success in mathematics, other domain-general factors are also likely to play a significant role, as well as other cognitive factors including attitude, motivation, and language ability (see Cragg & Gilmore, 2014 for a review). These skills are domain-general and are therefore not maths or literacy specific; rather they are skills that are important for learning and performance across all academic subjects, and they are likely to also influence science learning.

Many studies have considered the effect of executive function (EF) and language on academic achievement (Gathercole et al., 2003; Alloway et al., 2008) and there are some indications that working memory and EFs might well be implicated in science learning in preschool (Nayfield, Fuccillo & Greenfield, 2013). EFs are a group of processes that allow one to flexibly respond to environmental demands in goal-oriented, deliberate, and thoughtful actions (Miyake, Friedman, Emerson, Witzki, & Howerter, 2000). Studies have shown EFs begin emerging in infancy and continue to do so until late adolescence (Best & Miller, 2010; Huizingga, Dolan, & Van Der Molen, 2006). The three most commonly studied EFs are *inhibition* of automatic responses (supressing distracting information and unwanted responses), *shifting* of mental states (flexibly switching between different tasks), and *updating* or working memory (monitoring and manipulating information in mental space).

Studies have also shown that children have different profiles of performance across various components of maths knowledge and while they may have strengths in one area, they have weaknesses in others. For instance a study by Hecht et al (2002) found that while working memory related to fraction computation, it was not a predictor of conceptual understanding of fractions, hence the role of EF skills in the performance of mathematical calculations and how EF skills support the acquisition of new mathematics knowledge vary (Cragg & Gilmore, 2014). This suggests that it is entirely possible for children to have different competencies in relation to different science areas and that these science areas might be related to different EFs or domain-general skills.

Cragg and Gilmore (2014) stipulate that working memory has been shown to predict maths performance when using non-numerical stimuli as well as numerical stimuli (as shown by Gathercole et al., 2003). The significance of the non-numerical stimuli however, suggests that deeper domain-general factors are influencing maths proficiency as opposed to domain-specific numerical stimuli. There is some evidence that in 5-year-olds, executive functioning explains more variance in mathematics than in reading (Willoughby, Blaire, Wirth & Greenberg, 2012), while only inhibition and working memory skills predict English, maths, and science at age 11 (St Claire-Thomson & Gathercole, 2006) and age 14 (Nunes, Bryant, Barros, & Sylva, 2012), and shifting predicts maths and reading only throughout development (Yeniad, Malda, Mesman, van Ijzendoorn & Pieper, 2013). A study by Geary (2011) revealed how selective executive functioning skills are with regards to maths and reading: while the influence of working memory decreased with age for reading, it increased for maths. This suggests that although EFs are important for academic achievement, the exact relationships with different domains may vary.

Studies that have found working memory was not a predictor of conceptual understanding (Hecht, 2002; Hecht, Torgesen, Wagner & Rashotte, 2001) may explain why others have found relationships between science achievement and EFs but not with regards to conceptual development. Therefore research should take a componential approach to understand how EFs relate to science (Cragg & Gilmore, 2014). As the authors suggest, this necessitates further study into EFs beyond simple correlations and scores of academic achievement in order to examine the exact pathways involved because simply measuring performance at one time point does not reflect the learning of new science material. Indeed, research has regularly failed to account for science learning and achievement throughout primary school and as a result, little is known about the exact effects of domain-general capabilities on science learning.

Studies that have found positive influential effects of the executive functions on academic achievement for literacy and numeracy in primary school (Alloway, Gathercole, Willis & Adams, 2003; Alloway, Gathercole, Adams, Willis, Eaglen & Lamont, 2005; Alloway et al., 2008) and secondary school (Gathercole, Pickering, Knight & Stegmann, 2004; St Claire-Thompson & Gathercole, 2006) demonstrate that if literacy and numeracy are influenced by the same variables, it seems plausible that the same EFs would also be influential for science education.

4.2.2 Science and biology

Zaitchick and colleagues (2013) argues that EFs are likely to play a part in the development of biological theories. Opfer and colleagues (2012) also highlight the importance of domain-general features in science learning, particularly with regards to encoding new knowledge and subsequent ability to recall it. They argue that the gradual progression from novice to expert involves significant changes in what and how information is stored and retrieved from long-term memory, such as when solving problems. Science learning would require one to access long-term memory, and in that sense, the differences between novice and experts would also have strong implications for assessment and curricular design (Opfer et al., 2012). Studies of collaborative group work among primary aged children suggest that language may help to improve children's ability to manage information and aid categorisation thus reducing cognitive load by sharing information management (Tolmie, 2012).

More recently a study conducted by Zaitchick and collegues (2013) investigated the link between EFs and vitalist biology. Vitalist biology is the universal theory of ideas behind life, death, and health (Inagaki & Hatano, 2008). The researchers argued that the differences in the way in which children construct biological theories depends in part on the differences in their executive functioning abilities. They investigated this in a correlational study by giving children aged 5-7 a battery of EF tasks, and a battery of vitalist biology tasks. The former included the colour word test (testing working memory), and the hearts and flowers flanker task (Diamond, Barnett, Thomas & Munro, 2007), which tested working memory, set shifting, and inhibition. The battery of vitalist biology tasks included interviews based on Piaget's (1985) Animism interview, which probes what it means to be alive, the Death interview, probing what it means to be dead, and finally the Body Parts interview, probing the location and function of a series of body parts. Zaitchick et al (2013) created aggregate scores for EFs and for vitalist biology, and after conducting a regression analysis controlling for age and verbal IQ, came to the conclusion that EFs predict children's knowledge of vitalist biology.

However, while it is entirely likely that EFs do influence the development of scientific concepts, this particular study did not measure individual EFs with enough sensitivity to be able to explore exactly what executive functions (set-shifting, working memory, or inhibition) influence what aspects of biological knowledge. Moreover, the three vitalist constructs measured in this task were also measured separately but the scores for each summed to create a composite score, meaning that the task is inherently not allowing one to examine the links and influential pathways between concepts as well as within them. Results showed the three areas of vitalist biology were correlated with one another, which was expected given the fact that they all tap into similar areas of biology. But in order to test coherence, more in depth questions about all aspects of the three areas are required in order to observe how they influence each other. It might be individual executive functions are influential in different ways; hence a body of work to test specific general cognitive abilities with specific biological concepts is needed.

Cognitive flexibility is one EF that seems important in relation to science education given that children are often presented with multiple pieces of information/observations that conflict with their prior beliefs (e.g. Kallai & Reiner, 2010; Howe et al., 2012) and as such, children are likely to use this EF in order to modify or change a set of ideas given new and relevant information. Likewise, inhibitory control would be implicated in a similar way, so that children are able to discount any information that is accurate or indeed not relevant to their current representation about a particular scientific phenomenon. Lastly, the need to coordinate and integrate fragmented pieces of knowledge into a broader and possibly more coherent concept would implicate updating or working memory in children, also. Hence, this would suggest that the development of children's biological understanding would be heavily dependent upon the rate of development and sophistication of children's EF abilities. Note that this would also imply that any cases of impaired executive functioning would result in poor acquisition and conceptual development of scientific concepts.

The influence of EF of science knowledge would also suggest that children may struggle to coordinate and form explicit scientific ideas early on because these underlying domain-general functions are not fully developed. This might explain why children have many implicit ideas in relation to scientific phenomena but are unable to coordinate or explicate these rudimentary ideas until they are much older. In the context of biological concepts for instance, it may be that different aspects are initially fragmented, but start to cohere around age 7 or 8 as EF and attentional control become more sophisticated (Gathercole et al., 2004). From this stance, curricular sequencing might make sense, provided the timing of formal instruction is suitable.

Indeed Vosniadou (2014) investigated conceptual change processes and EFs and found significant correlations between the two. While it is certainly very likely that in order for domain-specific learning to occur there must be certain domain-general structures in place, the relationship between EFs and conceptual development within science education and in particular biological knowledge, has never been systematically researched. There is a desperate need for thorough investigation, which would naturally inform the relationship and influence of domain-general structures on domain-specific concepts.

4.2.3 Maths and science

Although there has been work to suggest number ability and science ability are connected in terms in of relative rates of progress (Gathercole, Pickering, Knight, & Stegman, 2004; St-Claire-Thompson & Gathercole, 2006), the exact nature of this relationship has not been examined in great detail either, and it is possible that the mediating role of language may also account for this (Tolmie, 2012). Studies have demonstrated the longitudinal predictive effect of short-term memory, working memory, and executive functioning in preschool children, on their academic achievement for mathematics by age 7 (Bull & Scerif, 2001; Bull, Espy & Wiebe, 2008), and on mathematical word problems (Zheng, Swanson & Marcoulides, 2011).

Some research has suggested that the relationships between mathematics and science knowledge are strong and that academic achievement in these areas is influenced by the same underlying EFs, and working memory in particular. This may be because both science and mathematics involve the manipulation of information, however with the latter, this manipulation of information is temporary given that its use is often in the service of problem solutions. This is also likely to be true for science, yet here it would seem more plausible that working memory would be especially important for the more permanent coordination of ideas, and conceptual integration over time.

It could also be that if working memory is to some extent used more on a temporary basis for mathematical problems, a child's mathematics ability could potentially be influential, if not predictive of children's scientific ability, given the importance of executive functioning and working memory for both. The perhaps more weak and temporary use of working memory in solving mathematical problems, and of course the success of the child's ability to do this, could in a sense influence their success to use working memory in a more strenuous exercise of manipulating, integrating, and coordinating conceptual ideas. This is speculation, however it may account for why academic success for numeracy seems to predict academic success for science also (Gathercole et al., 2004).

4.2.4 Inhibitory control in science

Inhibition is also a skill that is likely to be implicated in science learning. Zaitchick et al (2013) say the process of conceptual change are basically where the learner builds a new explanatory framework and inhibits previously useful (but inaccurate) frameworks for those same phenomena. Research (Shutlman & Valcacel, 2012) has already shown that naïve theories exist into adulthood. If one accepts the premise that children do indeed have naïve theories available to them after conceptual change or development, and that this persists onto adulthood (Shtulman & Valcarcel, 2012; Vosniadou, 2014) then this would require children to have good inhibition skills, which may increase with age.

Studies by Kwon and Lawson (2000) and St Claire-Thompson and Gathercole (2006) seem to suggest that inhibition, problem-solving ability, and spatial ability might be needed for scientific reasoning, however these same skills may not be needed for conceptual development, which may require less 'active' or intentional coordination of skills. For example, some studies (Mayer, Sodian, Koeber, & Schwippert, 2014) have shown that inhibition and scientific reasoning are not related. It could be that scientific reasoning is not the same as conceptual development, and the latter would require separate investigation.

Mayer et al (2014) argue that scientific reasoning is separable from intelligence and reading ability and designed a study to try and develop a paper-and-pencil scientific reasoning task for young children. In the study they developed materials to test domain-general scientific reasoning skills independently from earlier knowledge in specific scientific domains e.g.

physics, biology, and found there was no relationship between inhibition at age ten and scientific reasoning. It could be that the lack of relationship with inhibition reflects the lack of motivation for children to stop thinking about one thing, and instead think of something else, however it could also suggest that the test was not sensitive enough to measure inhibitory effects in scientific reasoning. With conceptual development, inhibition is likely to play more of an important role but the evidence from Mayer et al's (2014) study that it does not could be due to the fact that early knowledge is simply probabilistic learning and not theoretical. If it *were* theoretical then inhibition would most likely need to form an important part as inevitably theory change would require shifting thoughts some way or another in which case cognitive flexibility would be a skill that is also required, as mentioned earlier.

4.2.5 Working memory & systems thinking

As well the influences of EFs in general, there is reason to believe that working memory might be *particularly* important for science learning. For example, Vosniadou (2014) argues in order for children to hold a dual-representation of the Earth being spherical, and yet looking flat, working memory is likely to be required to maintain these multiple concepts. On the other hand, the fact that you have to maintain the Earth is spherical while inhibiting that it is not flat despite appearing to be may also require inhibition (Opfer et al., 2012).

Likewise, children often encounter many biological phenomena for which they may need to maintain multiple concepts or representations in their mind simultaneously. '*Systems* *thinking'* is the processes of thinking about how things influence one another within a system/environment. This may require a higher-order level of abstraction or cognitive capacity because it involves being able to think in more global terms about ecological systems and how various aspects influence each other and have long and short term consequences. Hipkins, Bull, and Joyce (2008) describe systems thinking with regards to ecology and note that the ability to hold multiple concepts in mind and notice the interactions does not occur until late childhood. They argue children's knowledge is very contextually based and they are not able to generalise their ideas across contexts. Perhaps then children's knowledge about their own cognitions is still developing, which could suggest that meta-cognition is a fundamental stage in conceptual development. However, we may also be considering the extent to which children are able to reach a level of abstraction when thinking more interactively, which is inevitably linked to the development of working memory in children at around age 8 (Alloway, Gathercole & Kirkwood, 2008; Kemp, Shafto, & Tenenbaum, 2012).

Working memory might also be implicated in a child's ability for diachronic thinking (Maurice-Neville & Montangero, 1992; Sander, Jelemenska, & Kattmann, 2006), which is the ability to think about biological processes over time, by understanding and representing key changes from the past, present, and future. A study by Maurice-Neville and Montangero (1992) investigated children's (aged 8-11) ability to reconstruct the evolution of forest disease by specifically examining the extent to which children introduced a temporal dimension to their reasoning. Children were shown a video about forest disease and then given pictures of a healthy spruce forest and a diseased pine tree. They were then asked to draw pictures of the spruce forest and subsequently a diseased pine tree, in the past and in the future and were interviewed about their knowledge and their drawings. Findings showed a distinct progression of the diachronic perspective with age, which the authors claim is slow and gradual until age 11 where the majority of children are able to think on a temporal axis. This ability may be particularly important for biological concepts such as evolution, where understanding natural selection over time is chief to understanding the entire concept of evolution. In fact thinking in a more global way might be the difference between novice and expert level knowledge. This is an interesting area to investigate given that has been a recent push in the new primary science curriculum in the UK towards practical skills and 'scientific thinking' and of course, the introduction of evolution as part of the new science curriculum.

4.3 Summary

The actual influences that EFs have on science learning, particularly biological learning have not been investigated in any great detail. A wealth of research from literacy and numeracy have suggested EFs play an important influential and in some cases predictive role of academic achievements in these subjects, and so there is a case for hypothesising an influential role of general cognitive abilities on science learning in primary school also. A systematic investigation is needed to assess how far any sources of shared variance between different conceptual areas are simply explicable in terms of general cognitive capabilities, because thus far studies have routinely failed to address this. Nonetheless, what is apparent is that there is no clear picture about conceptual development or progression in science, which can be used to steer curricular organisation. Thus a body of systematic cross-concept research that includes measures of general cognitive change is needed to clarify things. Studies have also suggested that it seems important to investigate language in order to access how it could be used as a driver for explicit thought, or conceptual development. Consequently, it becomes necessary to assess these variables longitudinally also, in order to track conceptual progression over time, which past studies have also failed to do. This would allow one to detect the developments and changes of children at both a group and individual level, and as such would also allow one to establish a sequence of conceptual development.

CHAPTER 5

5.1 Summary, rationale, and overview of current work

At the start of this programme of work it became very clear that there was reason to investigate how children are learning science in primary school given the recent curricular changes, but also the lack of progress in this endeavour in comparison to other core areas of the curriculum such as literacy and numeracy. More specifically important to investigate perhaps, is how children use the early concepts they acquire and how this might influence the acquisition of related or more complex concepts, and to understand the developmental trends in children's conceptual progression for ideas in naïve biology.

5.1.1 Conceptual change

Part of the reasons for why progress has been slow in understanding how children are learning science is the fact that conceptual progression for any discipline is under-researched. The debates around how knowledge is organised, whether in a theoretical framework or in fragmented pieces, have not reached consensus. These debates are important because understanding the ways children learn will inevitably have implications for curricular organisation and teaching practices. Understanding conceptual change is also important given research that has shown children have naïve biological concepts prior to any formal education, and that these concepts are often highly resistant to change (Driver et al., 1985). In order for instruction to be effective, one needs to establish what influences how children acquire naïve knowledge and how that knowledge might change and influence the acquisition of new and related knowledge over time. Understanding this, will aid the development of techniques or learning strategies that will provide learners with environments to foster accurate knowledge acquisition.

5.1.2 Essentialism

The research into naïve biology has also contributed to the debate around conceptual change, with the majority of theorists in favour of a theoretical framework. However, the essentialist paradigm that was used in this research seems to be highly flawed. As such, it seems unreasonable to continue to investigate children’s understanding of biological phenomena like inheritance, using a paradigm that constrains their answers, but more generally, also constrains progression in research. Up until now, much of the literature has been about confirming or disproving children’s notions of essentialism and by doing so, research has failed to consider other dimensions behind the inheritance concept and its relationship with other ideas. Instead it seems appropriate to start an investigation into children’s biological knowledge based on established research about what children *are* known to be capable of.

Firstly, it has been documented that children as young as six months have a strong tendency to detect the perceived relationships between events and to make probabilistic judgments about them, and even more evidence to suggest that this capability is generally established by around age 2 (Gopnik et al., 2001). Secondly, it has also been well documented that children have a natural tendency for categorisation with which meaningful language is also acquired (Gelman & Coley, 1991). What this indicates then, is that if the theoretical debates about inheritance and essentialism are put to one side, there seem to be two naturally occurring tendencies in children: perceptions of co-varying regularities, and categorisation.

Consequently, what is needed to move forward is to begin research with these already established ideas and project what conceptual development might look like if these were the point of departure, mapping observations against this projection to see how far it is borne out, using a more natural methodology that nevertheless still pushes at the boundaries of what children are capable of understanding or saying. Providing children with more familiar contexts is also likely to give them some assistance with which to ground their ideas, rather than using more obscure constructs as in the past (e.g. Gelman, 2003; Keil, 1989; Springer, 1999). Those earlier studies often asked children about unrealistic hypothetical instances, which did not necessarily allow children to answer questions in the context of what they knew, where it has been shown that knowledge is specific to that context in any case (Hipkins et al., 2008; Almeida, Vasconcelos, Strecht-Ribeiro, & Torres, 2013).

What might be helpful would be to assess children's knowledge of the same phenomena in multiple contexts which would allow one to infer firstly, whether their knowledge is context-specific and secondly, coherent (e.g. Almeida et al., 2013). If children are using probabilistic judgements and categorisation abilities to acquire information then it is likely that the first concepts to be acquired would be biodiversity concepts, which may lay the foundation for other logically related concepts. If so, we ought to understand, at least theoretically, how this might come about.

In this way more in-depth analysis about children's conceptualisations should help uncover the developmental trajectory of conceptual progression in biology, provided children across a relatively large age range are tested. Rather than focusing exclusively on inheritance, the natural extension would be to extend the range of concepts to include ecology, biodiversity, and evolution, since all these constructs share a number of important points of connection and overlap heavily as discussed in Chapter 2. These concepts in themselves are not unitary but are constructed of more important sub-elements and what is needed is a method of capturing children's understanding of all these elements in an even-handed fashion so that the relationships between them, and to more general cognitive capabilities, can be investigated.

5.1.3 General cognitive abilities

Research has suggested that EFs, aspects of memory, and inhibition in particular, might be key to understanding biological phenomena (section 4.2). Past research highlights that working memory and EFs may be predictive of children's understanding at preschool (Nayfield et al., 2013) primary school (Zaitchick et al., 2013), and secondary school (Gathercole et al., 2004; St Claire-Thomas & Gathercole, 2006). For these reasons it is hypothesised that general cognitive abilities would also have a large influence on conceptual progression in this study.

The links between general cognitive abilities and science learning have rarely been investigated for primary aged children, and where there is research of this kind, the focus has been on testing much more specific relationships, when ultimately a systematic investigation is needed to assess how far any sources of shared variance between different conceptual areas are simply explicable in terms of general cognitive capabilities. Any variance that is not explicable in terms of general cognitive abilities must imply conceptual integration and progression of some kind, which previous studies have routinely failed to address.

There has also been research to suggest that the numerical ability and language ability might also contribute towards children’s science knowledge. The exact relationships between language and number knowledge have not been assessed in any detail, and while there is work implicating numeracy and science knowledge are linked in particular (Gathercole et al., 2004), very little is known about how children’s early knowledge on general numeracy go on to influence scientific knowledge over time. Likewise the exact effects of language on conceptual change and progression have also never been addressed fully and given the importance of language to children’s categorisation, lexicalisation, and progress from implicit to explicit thought, it seems wise to assess its contribution towards learning biological phenomena in primary school, also.

5.1.4 Study rationale

Assessing naïve biology requires careful consideration given the fact that past research has often selected biological concepts arbitrarily. Inheritance is a key area that has always been a focus and it is also a concept that is formally taught in primary school. Given the past methodological flaws with early research, it seems sensible to assess children’s knowledge of inheritance using an unbiased and robust methodology to form a useful comparison with which to contribute towards the debate about conceptual development. Likewise, in order to assess progression of inheritance concepts, and examine the range of children's understanding on a variety of related concepts that are also part of the NC (DfE, 2014) for primary science: biodiversity, ecology, and evolution.

Inheritance, biodiversity, and ecology are topics that are currently taught in the NC in KS1 and KS2, which are logically related to inheritance concepts also and yet, have never been investigated thoroughly in the same context. Although evolution is a concept that was not covered by the previous NC (DfE, 2001), the current curriculum (DfE, 2014), includes evolutionary concepts as part of the KS2 primary science syllabus and so its investigation in this study is timely. These four constructs[2] (inheritance, biodiversity, ecology, and evolution) provide a logical set of ideas to explore the developmental path of concept formation based around inheritance, which has already been fairly well established in the literature.

Research into biodiversity, ecology and evolution reviewed in Chapter 2 also provide some key insights into what the developmental pathway of acquisition might be. The aims, research questions, and hypotheses are therefore discussed below.

[2] Note that henceforth 'constructs' refers to only inheritance, biodiversity, ecology, and evolution either collectively or individually.

5.2 Aims of the current research

The main aim of this current research is to examine conceptual development in a number of related biological concepts across primary school in effort to examine the range of children's understanding about naïve biological concepts and to comment on the developmental trajectory of these concepts in children aged 4-11. This research will also endeavour to develop a new methodology to assess biological knowledge in an unbiased and fully rounded manner, and to investigate the effects of general cognitive ability on biological understanding. By doing so, the outcome of this research should inform the current debates surrounding concept change and the structure of knowledge.

5.2.1 Research questions

Specific research questions pertaining to the aims of this research are:

1) Do children have sophisticated ideas about inheritance (assuming past research is correct)? If so, how do their ideas influence the development of concepts in other related biological constructs such as biodiversity, ecology, and evolution?
2) How are children progressing from having naïve and simple concepts to more sophisticated understanding of the same concepts? And how are related concepts influencing each other in their development and progression?
3) What is the influence (if any) of general cognitive abilities on children's biological understanding and conceptual development?

4) Does it seem as though children's concepts are theoretical or fragmented? And are there any developmental changes to this?

5.2.2 Hypotheses

The present study is exploratory and as such no concrete predictions can be made about the outcomes. To clarify, the purpose of the present research is not to discover what the end-point of knowledge might look like; rather it is to investigate the nature of concept formation and concept change, within biology in particular. Even though this work is exploratory, a review of the literature provides some indications as to the nature of conceptual progression. These are mapped out below along with predictions relating to the research questions outlined above.

1) Regarding the first two research questions, past research suggesting children's ideas about inheritance were coherent and theoretical is somewhat disputed by more recent research in the same area using more robust methodologies. Arguably children's knowledge about inheritance is still very fragmented even up to 14 years of age (Williams & Smith, 2010; Williams, 2012). Instead, children appear to be able to acquire a good understanding of biodiversity concepts which makes use of children's natural abilities to categorise observable information in their environment and make predictions based on this. Hence if this is the start-point of knowledge, it is likely that extensions of these initial ideas influence knowledge acquision in other related areas such as ecology (e.g. Assaraf & Orian, 2003; Sander et al., 2006; Almeida et al., 2013) where a context in

which to base ideas allows for more fine-tuned reasoning. It may then be that knowledge from these contexts feed into microevolutionary concepts, those that do not require a temporal axis to understand given that children find this harder (Shtulman, 2006). This ultimately results in children struggling to understand inheritance and macroevolutionary concepts such as natural selection which requires knowledge about inheritance without formal guided instruction. In this way it might be that early key concepts in one area may help to provide a basic foundation for other connected ideas to eventually form a coherent body of interconnected knowledge, which is compatible with curricular sequencing.

2) As for the influence of general cognitive ability on children's knowledge acquisition, it is likely that with increasing cognitive capacity to hold multiple concepts in mind simultaneously and coordinating these pieces of information would require working memory (Gathercole, Brown & Pickering, 2003; Alloway, Gathercole, Kirkwood & Elliot, 2008), hence one can predict that children with good working memory abilities are likely to have more advanced conceptual knowlegde. Similarly, the research literature presented in Chapter 4 also illustrated that children who had well-developed executive functions were also likely to be more advanced in conceptual knowledge (Cragg & Gilmore, 2014. This is due to the fact that they may be able to supress naïve theories (e.g. Shtulman & Valcacel, 2012) and demonstrate cognitive flexibility (Yeniad et al., 2013).

However, while it is the case that general cognitive abilities are likely to aid children in acquiring knowledge form their environment, their exact effects on conceptual *progression* are relatively unknown and in some cases, have not been shown to be predictive at all (e.g. Hecht et al., 2002; Hecht et al., 2001).

Nonetheless, given the influence of general cognitive abilities on literacy and mathematics (e.g. St-Claire-Thomson & Gathercole, 2006) their effects for science knowledge are also likely to be positive (Nayfield et al., 2013; Zaitchick et al., 2013).

3) Regarding research question 4, it may be that there is only a limited degree of co-ordination within particular areas and ultimately these ideas do not necessarily inform each other too well. The exposure children have to various species of animals through different means (e.g. school, zoo trips, media) and children's natural tendencies to perceive regularity and categorise may imply the construct of biodiversity emerges earliest. Indeed Hipkins et al (2008) demonstrated that children aged seven had clear understanding about biodiversity and basic ecological concepts such as interconnected food chains, presumably because this is a subject that can be taught relatively easily in contrast to ecological or evolutionary concepts. This would give rise to an experiential sequence which is likely to have cross-cultural variation yet mirrors the sequence of conceptual elaboration followed by Darwin's (1839, 1859) work: biodiversity, ecology, inheritance, and evolution. In this route of progression, the ideas may be somewhat isolated from each other in development and only inform each other at a basic level, at least to begin with. It also implies that children may have slightly different routes of progression depending on their individual experience.

On the other hand, conceptual development may take an entirely different route of progression, which remains context-specific throughout with the child piecing streams of information together rather like a mini-scientist as their cognitive capabilities permit (Gopnik, 1996). It could be that early biodiversity concepts develop first, as categorisation studies suggest (e.g. Gelman, 2015; Csibra & Shamsudheen, 2015), which feeds into rudimentary ideas around ecology and inheritance, which are specific to particular contexts. These may develop concurrently and eventually link individual variation to early evolutionary knowledge. In this account layers of knowledge are acquired gradually and thoughts are not integrated even at the level of topic. This route may also explain the fragmented picture research often depicts around children's early biological knowledge (e.g. diSessa, 1993; Williams & Smith, 2010; Tolmie, 2012). This account also implies that understanding may vary according to context (e.g. Hipkins, et al., 2008; Almeida et al., 2013), hence when seeking to examine conceptual elements held by children, doing so in contrasting contexts will help ascertain whether this account is valid. The implications of such a route suggests that children may have multiple epistemologies and that their knowledge is not theoretical as claimed by others (e.g. Gopnik & Wellman, 1994), but that knowledge acquisition and integration is a slow and complex process.

These potential routes will vary in the way that knowledge is acquired and used by children. As such each route would lead to contrasting developmental patterns in terms of both growth of understanding and the relationship between different elements.

5.3 Overview of Research

5.3.1 Overview of project design

A longitudinal project spanning two years will be presented. Primary school children from three cohorts were tested on their general cognitive abilities, and their biological knowledge. The same children were then followed-up one year later to participate in the same experiments as before in an effort to assess how their knowledge had changed over time. This enabled a developmental trajectory to be mapped out and informed on the differences in children's general cognitive abilities and biological knowledge across cohorts. It also informed on the specific influences on the development of biological knowledge in children over time.

5.3.2 Time line of data collection

A timeline beginning from October 2012 to October 2015 is shown below in Figure 5.1. The timeline illustrates the stages of methodological development, piloting, and subsequent two phases of data collection. Data collection was always in the autumn term of the academic year. This was particularly important to capture what children in the youngest cohorts knew upon school entry in September, and then what children across all cohorts knew at the start of the following academic year. There were also for more practical reasons for using these dates, such as the necessity to complete the entire project within a period of doctoral study.

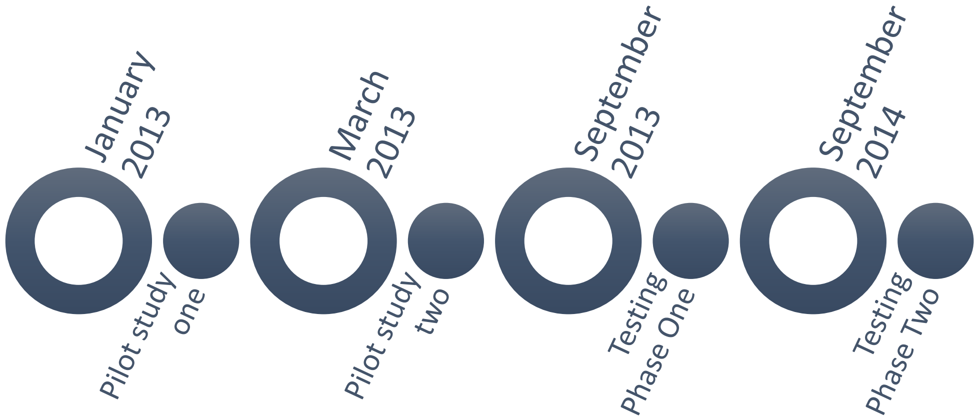


Figure 5.1. A timeline of data collection for this programme of study

5.3.3. Participants

The participants for the main study were recruited from three North London state schools. The sample for the pilot studies was recruited from an additional state school in North London, the final sample of which was 18. The first pilot study was carried out after the Christmas holidays in January 2013. Results from this study lead to further refinement of the methodology and a subsequent pilot study was conducted in March 2013. The second pilot study recruited a smaller sample of eight from the original sample of 18 from the same school. For both pilot studies, children were selected by classroom teachers from Year 1 (around age 5), Year 3 (around age 7), and Year 5 (around age 9) after an opt-out procedure for parental consent was used in agreement with the school.

As children could not be tracked from school-entry to the end of primary school, a triple-cohort longitudinal design was used for the main study. This meant three cohorts of children recruited in 2014 would be followed up in 2015, and as such, the study would be conducted in two phases. Testing Phase One was conducted in September 2013 for all three London schools consecutively. Children were recruited using an opt-in method of parental consent from Year Reception, Year 2, and Year 6. The final sample for Phase One was 138 across three schools. These same children were followed up approximately one year later in September 2014 for testing Phase Two. The final sample for Phase Two was 129 across the three schools.

The choice of age bands was made specifically to capture children's knowledge of biological phenomena at school entry age, during the middle of their primary school education, and towards the end of their primary school education. In this way key developmental patterns could be observed. Also, by tracking the sample over two years, a developmental picture for Reception, Year 1, Year 2, Year 3, Year 5, and Year 6 could be obtained. Note no data were available for Year 4 given the design of this study.

CHAPTER 6 - Methodological development

6.1 Overview

This chapter outlines the process of methodological development for a task assessing biological knowledge (accurate) and understanding (current beliefs, which may be incomplete or inaccurate). The previous chapters established the flaws of past methodologies focusing on children's essentialist beliefs. A key part of the research agenda for this study was to develop a methodology aimed at objectively examining children's knowledge of four related biological constructs. As such, this new method had to be piloted and refined in order to be used as a principled standard with which to measure children's biological knowledge. This is described below. Following the account of the development of this measure, the first pilot study is described. The results identified some changes that were necessary to improve the methodology. These are discussed before going on to describe and discuss the second and final pilot study.

6.2 Rationale

The aim of this programme of research was to track conceptual progression of biological constructs across primary school children aged 4 to 11 using a longitudinal design. Developing a methodology aimed at tracking conceptual development across a range of biological concepts has never been attempted in any real depth before.

The participant age groups were selected for specific reasons. Firstly there is a tendency in the literature to describe the key ages of conceptual development in the context of biology to be between the ages of 4-10 years (Carey, 1985; Piaget, 1970; Gelman, 2009; Vosniadou, 2014) yet no one study had examined conceptual change across this entire range. By doing this, a better overall developmental picture of conceptual change in biology was anticipated, which would allow tracking of any developmental trajectories and trends that would help to illustrate the patterns in conceptual development for biological phenomena during primary school years.

6.3 Development

A task used in a study of children's ecological understanding by Hipkins and colleagues (2008) was used as a basis for the development of this approach. In their study, children were provided with a two-dimensional drawing of an aquatic scene, which they were asked to complete with relevant drawings of other organisms, and then asked to explicitly write about the relationships organisms had with each other and their environment. Their study illustrated that the task was effective in getting children to think *overtly* about ecological systems and interdependence between the environment and organisms in that environment. Hence, they argue, children were actively trying to link related concepts together.

This task had the potential to be used for a combination of other biological concepts. However an issue with using the task was that it was specific to children aged above 8 years who were capable of writing about overt relationships. Hence the writing aspect was replaced using semi-structured interviews to allow children to articulate their thoughts through speech, based on a fixed set of criteria that would also allow the researcher to probe the limits of children's understanding. Also, a semi-structured interview format used in previous research (e.g. Williams & Smith, 2010; Williams, 2012) resulted in rich data about children's knowledge which previous language-sparse methods failed to capture (e.g. Gelman, 2003).

Investigating the effect of the mode of exposure for different contexts on children's biological knowledge was necessary because each context encountered presents its own challenges for knowledge acquisition and integration (Almeida et al., 2013), and because context has been shown to have an influential role in shaping the boundaries of conceptual development (diSessa, 1988; Hipkins et al; 2008, Vosniadou, 2014). Also depending on children's early experiences, different contexts may influence concept change and possibly help children explicate implicit thought (cf. Karmiloff-Smith, 1992). Using the study by Hipkins et al (2008) as a basis, children were presented with a selection of biological contexts within which they could draw relevant organisms. Drawing is something children naturally do from an early age and as the drawings were not coded themselves *per se*, they simply provided additional information to the answers children gave to various interview questions and acted as exemplar stimuli about which to ask those questions.

The contexts chosen were environmental contexts that children were most familiar with based on physical experience in urban or rural environments, and media or school coverage. Seeing the extent to which children are able to distinguish between these would provide useful information about how they are able to transfer knowledge across related contexts. A total of six environmental contexts for use in Pilot study 1 were developed (see Table 6.1).

Table 6.1. The six contextual scenes used in the pilot studies alongside the range of organisms

Context	Key criteria for selection	Animals used
Lake	Replication of Hipkins et al (2008) & is an obviously constrained environment with regards to the types of organisms one would expect to find there	Bird and Fish
Pond	More stereotypical looking than the lake and the body of water is less abstract than a lake, which might have been mistaken for the sea/ocean/river. The pond was a seemingly clearer example. Another familiar and constrained environment	Bird and Fish
Park	Familiar for urban/rural environments and constrained, but not aquatic	Bird and Worm
Field	As above (park), but less human-focused	Squirrel and Fox
Rainforest	Less familiar and loosely constrained, plenty of media/school coverage	Cheetah and Monkey
Savannah	As above, an alternative to rainforest	Lion and Zebra

The contexts were drawn by hand on A4 plain paper, and scanning these images on to a PC to colour and print. The actual content for each drawing was developed based on the example Hipkins et al (2008) used in their study. Copies of the contextual scenes that were developed can be seen in the appendix (A.1).

After developing the contexts, a set of interview questions was developed. The interview schedule for Pilot Study 1 can be seen in Table 6.2. As with Hipkins et al's (2008) study, questions about the two organisms on the page and their relationships with the environment were developed. Following this, children were asked environmental questions pertaining to ideas around ecology. Finally inheritance questions were included to examine whether findings would support previous work given that this method departed from the essentialist paradigm. These questions were organised logically and made to fit into the task context by asking children about the animals they would have drawn on the page, specifically about the kinds of offspring they would have and the types of traits that would be inherited, and why. It was thought that questions around inheritance, biodiversity, and ecology would allow the experimenter to probe children's answers about evolutionary phenomena sufficiently to grasp children's knowledge in these areas. Guidelines were established about the types of evolutionary concepts children might refer to (Table 6.2), which the experimenter could follow up on.

Table 6.2. Semi-structured interview schedule for Pilot study 1

Type	Question
Comprehension	What do you think this is?
	Why do you think it's a_____? How do you know it's a ______?
	Do you think these animals all live here? Probe.
	Have you been to/seen a ___before?
Drawings	
	Why did you draw a ___? (*understand which animals are relevant to the context and why)*
	Why did you draw a ____ next to/in/under/on the ___? (*understand that animals are suited to particular habitats)*

Inheritance	How did _____get born? (*understand the process of reproduction)*
	Children's answers to this question were probed.
Ecology	Why does the _____live here? (*understand animals are suited to the environment in which they live*)
	What would happen if the [lake dried up/there were no trees etc]? (*understand animals are dependent upon the environment in which they live*)
	Follow up
Biodiversity	Would there be as many types of animals around if [lake dried up/lions went away/no trees etc] *(understand interactive food chains/webs*)
	Follow up questions
Evolution	**Note there were no specific questions for Evolution but guidelines to assess children's knowledge*
	animals all compete for the same resources
	animals adapt to their environment
	change in the environment = adaptation
	consider why some animals are eating other types of animals

6.4 Pilot Study 1

Piloting for the first study was carried out between January and March 2013 to establish if the methods for investgating biological knowledge were suitable. The pilot study was carried out in a separate school to those eventually recruited in the main experiment.

A cross-sectional experimental design was used for this pilot study. The sample (N=18) for this pilot study was recruited from one North London state primary school. Six children were selected from Years 1, 3, and 5 to provide a concise cross-sectional overview of children's understanding at various ages.

The task was modelled on the task used in the study by Hipkins et al (2008). Children were handed a pencil to draw on a pair of randomly pre-selected contextual scenes from the set systematic pairings. A SONY DX3 digital voice recorder was used to audio-record the interview. Children were split into two groups: one group that completed the drawings before any questioning, and another group that was asked questions about the two organisms already on the page *before* drawing, and then asked the subsequent remaining interview questions. This was done so that any effect of questioning on children's drawings could be observed. There were a total of six contextual scenes used to show comparisons of performance across cohorts. Systematic pairings of each of the conceptual scenes were devised, and one of the pairs was randomly selected for each child. The pairs were: pond & park; pond & lake; pond & field; pond & rainforest; pond & savannah; park & lake; park & field; park & rainforest; park & savannah; lake & field; lake & savannah; lake & rainforest; field & savannah; field & rainforest; savannah & rainforest. The pairs of scenes were selected so that for every cohort, at least one of every *type* of context was used i.e. either a pond or lake; either a park/field; either a rainforest/savannah. The pairs of contexts were shuffled for each cohort and children were given the pair at the top of the pile at the start of the experiment. All ethical guidelines as stipulated by the British Psychological Society were followed.

The interviews with children took place individually in quiet meeting room in the school for effective recording. Children were presented with the first contextual scene and asked to identify it. They then partook in a general conversation about their experience of the context. From this point on, the procedure varied for those children in group one who completed the drawings before any questioning, and for group two who were asked questions about the two organisms already on the page *before* drawing, and then asked the subsequent remaining interview questions.

6.5 Findings-Pilot Study 1

Despite the weaknesses of curricular assumptions about conceptual progression, it is nevertheless the case that a serious attempt has been made to break different key areas down into definable components based on current science, and that this serves as a useful structure on which to focus, both in terms of obtaining an overall picture of primary children's understanding and how far it equates with curricular objectives at different ages. There is no current independent standard for measurement with which to calibrate conceptual development against. This is of course challenging but something that this study has attempted to tackle. The results from Pilot Study 1 made it clear that the interview schedule used there was not sensitive enough to detect the subtleties in children's knowledge about different biological areas. Thus the main issue with the previous interview schedule might have been that it was not grounded in any framework with which to systematically examine the four biological constructs, rather it was too abstract and vague. For these reasons it was felt necessary to ground the interview schedule in a comprehensive framework that would encompass all the sub-topics within the four biological constructs. An obvious way of doing this was to look at the NC for England and Wales. The curriculum has many aspects of the four biological constructs that have been broken down into smaller elements so that they can be taught across the primary years sequentially. For example, the curriculum fractionates science learning under broad headings such as *life processes and living things* (DfE, 2014), and based on earlier predefined terms, each element of knowledge under this banner was allocated to one of the four key biological constructs.

Therefore every element of the science curriculum (DfE, 2014) that related to any of the four biological constructs (inheritance, biodiversity, ecology, and evolution) was now taken as *core knowledge* (Spelke & Kinzler, 2000) for each biological construct, which when taken together allows, theoretically, for a full and coherent understanding of that broader construct. For example: "*recognise similarities and differences between themselves and others"* (DfE, 2014) was identified as an element of biodiversity knowledge as it emphasises the differences and similarities between and within species. To aid the children, many of the questions were constructed to ask about specific organisms and environmental artefacts in the previously selected contexts. Pilot Study 1 showed that the savannah and pond were the most successful contextual scenes; hence they were selected for a second pilot study.

Based on all the core knowledge elements identified, interview questions were developed (Table 7.4 in Chapter 7). Note that a number of questions addressed more than one core knowledge structure and not necessarily within the same overarching construct, where this is the case, the question number is cross-referenced.

A final point to note is that not all the questions developed were context-specific, particularly those concepts that were centred around humans and on inheritance concepts, because the results from the previous pilot study had shown children had non context-specific ideas about. Hence it was decided that any question children had generic knowledge on (often human-based questions that wold not need to be repeated) would be excluded for the second interview i.e. children would answer all questions for the context they received (e.g. savannah), and to save on testing time the generic questions from the same interview schedule would be removed for the second context (e.g. lake). The generic questions that were only asked once are highlighted in bold text in Table 7.4 in Chapter 7.

Pilot Study 1 also revealed that there were no obvious differences between the group that was assigned drawings prior to the interview, and the group that was told to draw *during* the interview; hence the groups children were in did not seem to make a difference in the expression of children's knowledge about biological constructs.

With regards to the drawings, many of the children became preoccupied that their drawing skills were being assessed despite reassurance that they were not, and so were very reluctant to draw anything at all. Those that did draw often drew things that they overtly expressed as not being relevant but drew anyway. In this respect it was felt that in further developments of this methodology, the drawing aspect should not be included.

6.6 Pilot Study 2

The format for this study was essentially the same as the first, though testing time was substantially reduced to approximately 20 minutes. They were shown both contexts in a counterbalanced order and asked comprehension questions, followed by questions relating to the four biological constructs. The interview schedule was fixed and was exactly the same for both contexts but depending on which context the child received first, the interview for the second context was administered with the removal of the generic questions as described above. The interviews were audio-recorded and transcribed for later analysis.

Six children were randomly chosen from the original sample of 18 in Pilot Study 1 on the basis that they had not previously seen either the savannah or pond contexts. There were 2 children from each cohort. This pilot study used similar materials as the last one, two A4 sheets of paper with drawings of four animals commonly associated with the scene. A SONY DBX3 voice recorder was used to record the interviews as before.

Children were presented with one contextual scene after the other (counterbalanced) and were briefly asked to identify the animals and the place in the picture. Following this the children were interviewed using the schedule developed on the basis of core knowledge (see Table 7.4 in Chapter 7).

6.7 Findings-Pilot Study 2

The developments made to the biological task for Pilot Study 2 following the results from Pilot Study 1 significantly improved the task. The results from Pilot Study 1 had shown that there was a need for more specific and detailed questions that addressed all aspects, or elements, pertaining to each biological construct. This allowed one to capture the full range of knowledge and the level of understanding (accurate or otherwise) children might have across and within each biological construct. Regarding the contextual scenes, all children were able to provide good knowledge about each context and were familiar with the environments. The nature of the interview also made administration of the task much easier.

A content analysis (Krippendorff, 1980) was used to ascertain general patterns and themes in children's findings. Generally it seemed as though children's early ideas about inheritance were fairly fragmented. Generally there did not seem to be evidence of coherent ideas about any biological concept, even among the older children as their knowledge was often fragmented showing no evidence of being theoretical. Previous work has depicted inheritance as the starting point for biological conceptual development, yet this pilot study suggested biodiversity might be instead, as this was the area children seemed to have the most knowledge about, even from a very early age.

As expected, ecological and evolutionary constructs appeared to be much less understood. If true, this could possibly be due to a lack of diachronic thinking ability (Maurice-Neville & Montangero, 1992) as principally, these constructs need to be understood with regards to change across a temporal axis. This can only be explored after the main study.

By systematically and comprehensively breaking down elements of core knowledge that together would form a coherent understanding of a biological construct, the current task allowed examination about which sub-areas in the contexts children might have more knowledge on, but also allow one to be confident in the fact that all aspects of each biological construct had been assessed. This is something that previous tasks using the essentialist paradigm had failed to do. Moreover, past studies had failed to examine children's understanding about a range of related biological phenomena, which the current biological task developed here allowed.

Overall, the task developed and described above was considered fit for purpose but given the results from Pilot Study 2, it was decided that a few minor adjustments should be made.

The two original contexts of the savannah and lake were kept the same, but were transferred onto A3 sized sheets of paper which were laminated. This was to ensure that the task could be used repeatedly without damage. However, during the second pilot study, children often used their hands or objects to obscure certain aspects of the picture in order to better explain themselves when asked a particular question, so it was felt that by having a more interactive picture that children could manipulate, they would be better equipped to articulate themselves. The materials were the only aspect of the task that had changed, the interview schedule itself was kept the same. For these reasons a third pilot study was felt superfluous.

CHAPTER 7

7.1 Overview

The following chapter describes the methods used to obtain data for this thesis. Data collection was carried out over two years throughout two phases (see Figure 5.1 in Chapter 5 for a timeline of data collection). This chapter begins by providing an outline of the design of the study, highlighting the main measures used for data collection. It then goes on to describe the participants and ethical considerations connected to the study, and finally the materials and procedures used in the study, both in Phase One and Phase Two.

7.2 Design

A triple-cohort longitudinal design was used for this research. Children across the entire primary age range (age 4-10 years) were recruited in three cohorts: at school entry, halfway through primary education, and towards the end of primary education. The children in these cohorts were then followed up approximately one year later (Phase Two) so that a brief picture of development could be observed across primary school.

The first phase of data collection was conducted in the Autumn/Winter term of 2013, and the second phase conducted in the Autumn/Winter term of 2014. In order to reduce testing time, there were two separate testing sessions during each Phase: one session to administer the general cognitive measures that included tests on working memory, short-term verbal recall, and visuo-spatial memory. There were also tests on semantic inhibitory control, cognitive flexibility, receptive vocabulary, and number knowledge. As discussed in Chapter 4, the measures of general cognitive abilities were included due to past studies highlighting a link between memory and EFs to areas of children's scientific knowledge at preschool (Nayfield et al., 2013), primary school (Zaitchick et al., 2013), and secondary school (Gathercole et al., 2004; St Claire-Thomas & Gathercole, 2006). For these reasons the influence of different aspects of memory including short-term verbal, visuo-spatial, and working memory, and the three core EFs as described in section 4.2.1, (cognitive flexibility, inhibitory control, and updating) were assessed.

A second testing session approximately three weeks later was conducted to administer the biological task (described in Chapter 6) for measures of children's level of knowledge on inheritance, biodiversity, ecology, and evolution.

Phase Two of data collection was exactly the same as Phase One, with the exception of an additional measure of expressive language included in the second session of Phase Two only. In total, children had four testing sessions across two years, lasting approximately 30-40 minutes per session.

7.3 Participants

7.3.1 Ethical approval

Ethical approval for this study was obtained from the Faculty Research Ethics Committee for the Institute of Education via the Department of Psychology and Human Development, who decided that there were no known risks associated with this study. The Disclosure and Barring Service carried out enhanced checks prior to any testing. Written consent was obtained from all parents/guardians of children participating in this longitudinal study, and verbal consent was always acquired from each child prior to any testing. The main experimental procedures were outlined to the parents via the consent forms and were explained to the children before each testing session. Children (and their parents) were made aware that their participation was entirely voluntary, was not linked to any academic assessment, and that they were free to withdraw at any point. Children and their parents were also made aware that all data collected would remain anonymous and strictly confidential. At the end of data collection, children were fully debriefed about the aims of the study and given the opportunity to ask questions. Participants in the study were not deceived in any way and there was no obvious risk of physical harm or psychological distress.

7.3.2 Selection of schools

Participants were recruited from three schools in London, United Kingdom. Each was a state school based in a deprived area of London and had a similar profile of intake, in that nearly half of all pupils were from ethnic minorities, and of these, approximately half spoke English as a second language.

Recruitment of the schools was done opportunistically, schools were contacted about the project and all agreed to take part in the study because they shared a key interest with the project in ultimately trying to improve the standard of science education. As the new science curriculum was due to come into place the following year (2014), many teachers in the schools expressed their interest in finding out the results of the study, particularly with regards to children's knowledge about evolutionary concepts. Also, given the length of commitment to the project, one of the schools agreed to participate in exchange for a brief teacher training session on new developments in science education research, while the others agreed to take part in exchange for a certificate of research participation.

7.3.3 Selection of participants

This study recruited children for the longitudinal study in three cohorts prior to testing in 2013. Children from Reception (age 4/5 years), Year 2 (age 6/7), and Year 5 (age 9/10) were recruited. These same children also took part in Phase Two of the study where they were subsequently in Year 1 (age 5/6), Year 3 (age 7/8), and Year 6 (age 10/11) respectively, as shown in Table 7.1.

Table 7.1. Age bands of participants in the three cohorts with corresponding Year at school across the two phases of testing

Phase	Age	Year at School
Phase One (Autumn/Winter term 2013)	4/5	Reception
	6/7	Year 2
	9/10	Year 5
Phase Two (Autumn/Winter term 2014)	5/6	Year 1
	7/8	Year 3
	10/11	Year 6

In agreement with the schools, invitation letters and consent forms were sent out to parents of pupils in the desired age range, hence sampling was purposive given the selectivity. Class teachers were asked to exclude children with developmental disorders or where speaking English language may have been an issue, because of the nature of the study protocol. The resulting sample composed of 138 children (73 females) divided into three cohorts. These children participated in both phases of the main study.

7.3.3.1 Participants in Phase One (2013)

In Phase One, there were 46 participants from school OG, 47 participants from school GT and 45 participants from school BP (see Table 7.2). Of the total sample, 48 children were in Reception (mean age = 56.96 months, SD = 3.50), 45 from Year 2 (mean age = 81.13 months, SD = 7.12) and 45 from Year 5 (mean age = 115.29 months, SD = 3.68). The difference in standard deviation for Year 2 children suggests a larger distribution of age in months from the mean in comparison to the other two cohorts where the standard deviation is lower. There was also a significant association between year group and school as is reported in the next chapter.

Opt-in letters were sent home to parents/guardians and informed consent was obtained from all parents/guardians of children in the study prior to testing. The letters that were sent home included information about the procedure and rationale behind the study. Parents were told that given the new changes to the primary science curriculum due to be implemented in September 2014, a study was being conducted to investigate children's understanding of an area of science in an effort to examine the efficacy of the curriculum. The letter also highlighted that science is a difficult area for many children and that the study aimed to understand exactly how children were learning certain scientific concepts in an attempt to try and improve teaching practices (appendix A.3).

Attached to the parental consent letters was also a brief questionnaire about demographic information, described in section 7.4.1. Out of the sample of 138 in Phase One, 116 parent demographic questionnaires were returned (84.06% response rate across the whole sample). One child withdrew at the start of the study and although parental consent was obtained, the child was very reluctant and so was not pursued.

Table 7.2. Distribution of participants in each cohort relative to the school they were recruited from

	School attended			
	OG	GT	BP	Total
Reception	13	16	19	48
Year 2	14	9	22	45
Year 5	19	22	4	45
Total	46	47	45	138

7.4 Materials, administration & scoring

A combination of standardised tests taken from a selection of test batteries was used for measures of cognitive indices. Measures of working memory (digit recall, block recall, and backwards digit recall subtests) were taken from the Working Memory Test Battery for children (WMTB-C; Pickering & Gathercole, 2001). The British Picture Vocabulary Scale (BPVS-3; Dunn & Dunn, 2009) was used as a measure of receptive language, alongside the Number Knowledge test (NKT; Okamato & Case, 1996) as a measure of number knowledge awareness. The Wisconsin Card Sorting Task (WCST; Grant & Burt, 1948) was used as a measure of attention shifting/cognitive flexibility, and finally as a measure of semantic inhibitory control, an adapted version of the Chimeric Animals Stroop task (Wright, Waterman, Prescott, & Murdoch-Eaton, 2003) was used.

7.4.1 Parent Questionnaires

The parental questionnaire included items such as: number of adults in the home, number of younger siblings, number of older siblings, socio-economic status (SES), preschool attendance, languages spoken at home (English only, other language only, mixed bilingual), mother's education level, father's education level, mother's occupation level, and father's occupation level. Parents/guardians were informed that their data would remain anonymous, confidential, and that they were free to withdraw their child and their data from the study at any time without question.

It was felt necessary to measure the language status of the child on the grounds of language being a potential mechanism for conceptual change (section 2.5). There is also reason to believe that verbal competency may be influenced by the number of adults and children in the home (Hart & Risley, 1995). Additionally, past studies have established strong links between levels of parent education and academic achievement of children (Alexander et al., 1993; Duncan et al., 1994), and there is equally strong research that links this to SES (see Bradley & Corwyn, 2002 for a review). Finally pre-school attendance was measured as this has often been associated with later academic achievement (Rashid et al., 2013).

SES was measured by whether or not the child received free school meals, and the score was therefore binary. Data for the language(s) spoken by the child were classified into three groups (monolingual English, monolingual other language, bilingual including English) and each of these groups was coded as 0, 1, 2 respectively. Mother and father's education level was coded on the same scale: 0 for no education past high school, 1 for A Level or equivalent, 2 for undergraduate degree or equivalent, 3 for Master's degree or equivalent postgraduate training, and 4 for PhD, equivalent training, or multiple degrees. Mother and Father's occupation level was also coded on the same scale: 0 for unemployed or retired, 1 for unskilled manual labour, 2 for skilled manual labour, 3 for skilled labour not manual and professional office workers, 4 for training required, 5 postgraduate training required, professional level occupation.

7.4.2 Digit Recall

This subtest was taken from the WMTB-C (Pickering & Gathercole, 2001) as a measure of verbal short-term memory. In this task, the experimenter reads out a sequence of three numbers one after the other, which the child has to repeat in exactly the same order. "*I'm going to say some numbers. I want you to listen carefully, and say the numbers back to me in exactly the same order that I did, OK? Let's practice.*" The child is then given three practice trials of three digits. If the child could not do this accurately, trials began using a sequence of two digits. If the child was successful in the practice trials, they began the task using three digits. "*OK great. Let's start. Remember, say the same numbers after I've said them in exactly the same order that I did."* If the child was able to do this correctly at least four/seven trials in a block, they moved onto the next block where an extra digit was added on to the random sequence of numbers until the child failed to recall the correct sequence at least four times in a given block. The maximum number of digits a child was able to successfully recall was taken as their *digit span*. Each trial a child got correct was scored as one. The number of correct trials before digit span was reached was taken as a measure of verbal short-term memory.

7.4.3 Backwards digit recall

This subtest, also taken from the WMTB-C (Pickering & Gathercole, 2001), is a commonly used measure of working memory. The task works in a similar way to digit span in that the experimenter read out a list of two digits to begin with, which the child had to repeat in the correct backwards sequence. "*I'm going to say some numbers. I'd like you to listen carefully and say the same numbers that I said, but backwards. For example if I say '1…2' I want you to say '2…1' OK? Let's practice."* The child was then given three practice trials with two numbers. If they were able to do this correctly they moved on to the task.

For younger children aged 4-5 years, and for those who failed the practice trial of two digits, a number line labelled from 1-10 was shown and two cards with random numbers from 1-10 were selected by the experimenter and placed in front of the child. These children were given the same instructions as above, but the experimenter pointed to the cards in the forward and backwards direction to illustrate how the child must respond correctly. This was done for all of the practice trials as specified in the WMTB-C manual. Following successful practice trials, the number line and cards were removed and the task began with a sequence of two digits. "*OK great. Let's start. Remember, say the same numbers after I've said them in the backwards order like we practiced."* Each trial a child got correct was scored as one. If the child was able to do this correctly at least four/seven times, an extra digit was added onto the random sequence of numbers until the child failed to recall the correct sequence of numbers at least four times in a given block. The maximum number of digits a child was able to correctly recall in a backwards sequence was taken as their *backwards* *digit span* and the correct number of trials before backwards digit span was reached was taken as a measure of working memory.

7.4.4 Block recall

This subtest taken from the WMTB-C (Pickering & Gathercole, 2001) is a measure of visuo-spatial memory. It is essentially a non-verbal version of the digit span task whereby the experimenter laid out a tray with a number of raised blocks on that were arranged randomly. These blocks had no distinguishing features and were only numbered on the side of the blocks visible to the experimenter. The experimenter tapped a sequence of three blocks with their index finger. The child then had to recall the sequence of the blocks that were tapped, and do the same. "*Have a look at these blocks. I'm going to tap some blocks with my finger like this*, *when I'm done, I want you to tap exactly the same blocks that I did in exactly the same order, OK? Let's practice."* The child got three practice trials with a sequence of three blocks. If a child failed to do these correctly they began the task with a sequence of two blocks. If they passed the practice trials they began with a sequence of three blocks. "*OK great. Let's start. Remember, tap the same blocks after I've tapped them in exactly the same order that I did."* Each successful trial received a score of one. If the child got at least four/seven trials correct, they moved onto the next set of trials where an extra block was added onto the random sequence until the child failed to recall the correct sequence of blocks at least four times out of seven trials. The maximum number of blocks a child was able to successfully recall was taken as their *block span* and the correct number of trials before block span was reached was taken as a measure of visuo-spatial memory.

7.4.5 Receptive language

The BPVS-3 (Dunn & Dunn, 2009) was used a measure of receptive language. This task involved showing children a page of four pictures that were numbered. The experimenter then said a word and the child was asked to point to the picture, or say the number of the picture that they thought was the word the experimenter had said. "*Have a look at these four pictures. They each have a number at the bottom, 1,2,3,4* [points to the numbers] *I'm going to say a word. I want you to listen carefully and tell me the number of the picture, or point to the picture, that you think is the word that I just said, OK? Let's practice."* Children were given two practice trials, the selection of which was based on their age group as Dunn and Dunn (2009) specify. In the unlikely event that children failed the practice trials, the task was explained again and the experimenter pointed to the correct picture by saying "*I said* [ball], *can you see this is a picture of a* [ball]? *I want you to point to the picture that is the word that I said*" Depending on the age of the child, the task was started at different blocks as specified in the BPVS manual. "*OK great, let's start. Listen carefully to the words I'm about to say*." The experimenter showed children a new set of four pictures per page prior to saying a new word. Children were scored either zero or one for an incorrect or correct trial. There were twelve trials in a block and the trials became progressively harder in each block. If a child answered eight or more trials incorrectly in a block moving forward, the task was stopped at the end of that block. If in the very first block of trials children scored more than two trials incorrectly, they went back towards the previous block until a baseline was reached. The baseline was established if children got either one or two items incorrect out of a block of twelve. The number of correct trials was taken as a raw score for receptive language.

7.4.6 Number Knowledge

The NKT (Okamato & Case, 1996) was used as a measure of number knowledge awareness. This task asked age-specific questions to children and depending on the age of the child the task began and ended at certain points. "*I'm going to ask you some questions about numbers now. Remember this isn't a test, if you're not sure about the answer that's OK, just try your best."* Children were asked questions relating to number order and simple arithmetic following the list of questions provided in the NKT such as "*what number comes two numbers before 5?"* This got progressively harder "*what number is 9 numbers after 999?"* Children were scored one for every question they get correct in an age-appropriate block. If children got five consecutive correct answers, they moved onto the next block. If children failed to score at least five in their age-appropriate block, they were asked questions from the previous block for younger children until a baseline was reached. The correct scores relative to the number of questions administered created a percentage overall score which was used as a measure of number knowledge.

7.4.7 Semantic Inhibitory control

The chimeric animals task was converted from paper task to a digital version presented on a MacBook Air 11.6 laptop. This task is described in more detail below given that it was adapted from the original study. The original version of this task can be presented on either Eprime software or using paper cards. Children were shown four cartoon-illustrated animals: cow, sheep, duck, and pig in the congruent trials of the task (see Figure 7.1). In the incongruent trials, children were presented with the same animals, but the animals' heads and bodies were mixed up e.g. a pig's body with a duck's head. In the incongruent trials, children had to inhibit their natural tendency to name the animal (under speeded conditions) by looking at its head, and instead focus their attention on the body of the animal to get a correct answer. Children were shown a total of eight trials in a random order per block and were given a short break at the end of each block. They were shown two incongruent blocks, and six congruent blocks. Children were scored on the number of errors in each block of trials, and the time taken to complete each block.

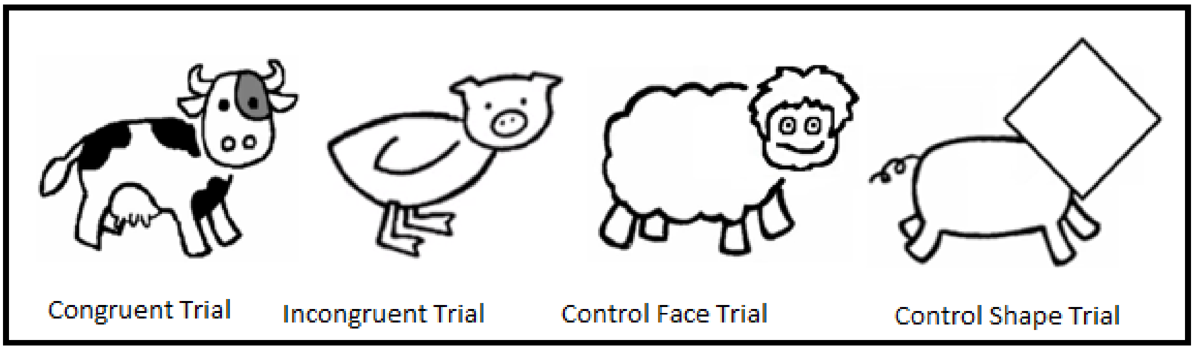


Figure 7.1. Examples of the four types of trial used in the Chimeric Animals Stroop task taken from Wright et al (2003) with permission

As suggested by the original paper (Wright et al., 2003), other blocks looking at children's attention to the faces were also included. The first of these blocks of trials was known as 'control-faces' where the body of the animal had a human (cartoon) head on it instead of an animal and so children once again had to focus on the body of the animal, this time without the distraction of another animals' head. The second of these blocks of trials was known as 'control-shape' where no face of any kind was used on top of the animals head and instead a selection of shapes (e.g. triangle, circle, square) were used to replace the animals head, once again focusing the child's attention on the animal's body without the distraction of any type of face shape at all, animal or otherwise. Therefore, there were four types of blocks of trials: congruent, incongruent, control-face, and control-shape. Each block had eight trials, and the incongruent, control-face, and control-shape blocks were always separated by a congruent block of trials. The final sequence was: congruent, incongruent, congruent, control-face, congruent, control-shape. This sequence was repeated twice, using a different order of images within each trial. The images were standardised so that each animal was paired with each shape, face, or head and body of another animal. As there were three blocks of congruent trials in the final sequence, two blocks were the same and so had to be counterbalanced for example: block A congruent, block A incongruent, block B congruent, block A control face, block A congruent, block A control shape. These created two sets of trials, set A (where there were two congruent A blocks, one congruent B blocks, and the remainder of the other blocks were all block A) and set B (where there were two congruent B blocks, one congruent A block, and the remainder of the trials were all block B). These were counterbalanced during administration with half the sample of children starting with set A then set B, and the other half starting with set B then set A.

To begin with, children were shown pictures of a cow, duck, sheep, and a pig to ensure they were able to participate in this task and following successful labelling of these animals, they were then instructed to name the image on the screen as fast as they could. If children were unable to label the animals correctly, the researcher went through the label for each animal and repeated the trial. If the child was still unable to label the animals correctly, they were excluded from this particular test. After this, they were given a practice trial using images of vehicles taken from Wright and colleagues (2003) with permission. When a child uttered the first syllable of a word, the experimenter immediately switched to the next image to keep the trials as fast as possible. Children were instructed not to use words such as “the” or “a” before naming the image, and if they did they were reminded of this before the main trial. After the practice trial children were once again reminded of the instructions and were asked to “*name the animal on the screen as fast as you can. If you make a mistake, don’t worry, we’ll just move onto the next picture, OK?”* The instructions were also written on the screen in front of them. When the child was ready, the congruent trials were administered, and the experimenter timed the entire block of trials from the first utterance the child gave for the first picture to the first utterance the child gave to the last picture. No feedback was given throughout the trials, and if a child made a mistake the experimenter continued on to the next picture. If children took longer than 5 seconds to name the animal, the experimenter moved on to the next picture as in the original study (Wright et al., 2003).

After the congruent trials, children were then told that they were about to see some funny pictures where the head of the animal was different to its body and were shown a practice picture (again taken from Wright et al 2003, with permission) for which they were asked to name the body. When the children were able to understand the task, they were told the instructions of the task again, which were also displayed on screen “*This time I want you to name the body of the mixed-up animal as fast as you can. If you make a mistake, don’t worry, we’ll just move onto the next picture, OK?”* The incongruent trials were presented when the child was ready. These trials were followed by the congruent block again, where the child was given the same instructions as before, followed by the control-face block (“*this time you’re going to see some funny animals again with silly cartoon faces, I want you to just name the body as fast as you can ok? Don’t worry if you make a mistake, we’ll move onto the next picture.*”), the congruent block, and finally the control-shape block (“*this time* *you're going to see some funny animals again with shapes covering their faces, I want you to just name the body as fast as you can ok? Don't worry if you make a mistake, we'll just move onto the next picture, OK?).* This concluded the administration of set A of the trials. Set B was conducted a while later so as to maintain the child's attention of the task in the same way as set A described above, with the omission of the practice trials. After the end of each block of eight trials (16 in total from set A and B) the number of errors for each block was also recorded. Hence, the measures obtained from this task were: number of congruent errors for set A, set B, and total; number of incongruent errors for set A, set B, and total; number of control-face errors for set A, set B, and total; and number of control-shape errors for set A, set B, and total.

7.4.8 Cognitive flexibility

The bulk of the WCST remained the same as the original administration, but was slightly adapted. Children were presented with 4 target cards as in the original, a red triangle, two green stars, three yellow crosses, and four blue circles. Children were then given a variety of cards that differed from the original set either by the number of shapes on the card (e.g. two blue circles), the colour of the shapes (e.g. one green triangle), or the shape on the card itself (e.g. four blue crosses). All the cards were created to differ from the target cards either by shape, colour, or number and were repeated twice; hence four colours, four shapes, and four number variations leads to (4^3 = 64)2 = 120 cards in total were created. Children were told they were going to play either the shape game, the colour game, or the number game, where the aim of the game was to sort out the cards into matching piles underneath the target cards by shape, colour or number respectively. After three practice trials to make sure children understood each game, they were then told "*now we are going to play a guessing game because I'm not going to tell you how I want you to sort the cards, I want you to guess. Is it the colour game, the shape game or the number game? OK? But I might also change the game as we play, so we might start off playing the colour game but I might change my mind and want to play the number game instead. Your job is to try and guess what game I'm playing with you."* Children were then told a correct answer would get a "*yes*" response, and an incorrect answer would get a "*no*" response. The cards were presented to children in a random sequence after shuffling. The experimenter always made switches to the rule after 10 cards had passed. The sequence of the rules per 60 trials remained the same at: Shape, Colour, Number, Colour, Number, and Shape. Children were scored on the number of rule changes they were able to detect. This was measured by placing the card in the correct target box 3 times in a row. If a child was able to do this, they were scored as having understood the rule. Children were able to get a maximum score of six per 60 trials, and 12 in total after 120 trials. Children were also measured on the number of perseverative errors they made. This was measured by counting the number of times children kept on sorting with the previous card rule consecutively, after a switch in the rule had taken place. If children sorted using a different rule from the previous one or the current one and still made an error, this was not counted as a perseverative error. Children were able to score a maximum of 50 perseverative errors per 60 trials.

It was thought that presenting children with all 120 trials of this task in one go would cause boredom or fatigue, and possibly reduce attention. For this reason the task was divided into two halves of 60 trials, presented at the beginning and at the end of the testing session.

Hence the final sequence of testing for session one was: WCST set one, Stroop task set A, digit recall, block recall, backwards digit recall, BPVS, number knowledge, Stroop task set B, and WCST set two. Note set A and B for the Stroop task were counterbalanced.

7.4.9 Biological task

Two A3 laminated colour scenes of a savannah and a lake were used following development from Pilot Study 2. Cut-outs of laminated animals and plants were also included for children to place on the scene and manipulate. Figures 7.2 and 7.3 illustrate the two contextual scenes (savannah and lake).

Children were tested on their biological knowledge on a separate occasion approximately three weeks after the cognitive indices testing. Children were randomly assigned a counterbalanced order of contextual scenes prior to testing. Testing was conducted with each child individually in a quiet room. The materials were set out on a table and the child and the experimenter were sat at the table next to each other, with the child always seated to the experimenter's left. Average testing time per child was approximately 30 minutes.

Upon entering the child was reminded about the background of the study, that the experimenter was interested in how children were learning science, and that today they would be asked about some things related to science learning. They were told that there was no right or wrong answer and that the experimenter was not testing them, they were just interested in what the child thought about some of these things. The child was informed about what the biological task would involve: "I*'m going to show you two pictures* *one after the other. Let's do this one first*, *I'm going to ask you some questions about the things that you can see on this picture ok?"* They were also made aware that the conversation would be recorded and that the experiment was not an academic test, their interview would remain confidential and anonymous and that they were free to withdraw from the study at any time. A SONY DBX3 digital recording device was used to record the interviews. Upon gaining the child's verbal consent to record the interviews and participate in the task, the digital recorder was switched on and the child was asked about their familiarity with the place in the context, and their experience or knowledge about it, if any. Children were asked to look at the picture in front of them: "*Have a look at this picture, what do you think this is a picture of? Have you been to a place like this before?"* Children were given the opportunity to name the place in front of them. If they named the place incorrectly, they were told the correct name of the place, and asked if they had been here before, and asked to tell the experimenter about their experience. They were then asked to place the animals and scenery on to the picture with the help of the experimenter. "*That looks great! Now I'm going to ask you some questions about what you can see in front of you. If you're not sure about the answer that's ok, I'm just interested in what you think about some of these things."* If the child had any questions, the experimenter answered these.

The interview was then conducted using the predetermined set of questions (Table 7.4) that were developed and tested during two pilot studies described in the previous chapter. These questions were developed based on the core knowledge structures extracted from the NC. The interview questions were always in front of the experimenter, and depending on the answers of the children, these were then probed, repeated if the child misunderstood, or omitted if the child had already answered them as part of another question.

As children were interviewed for both contexts consecutively, there were some questions that previous piloting had shown were too repetitive to be asked twice. These were questions mainly around humans and questions about factual knowledge, where children's answers were unlikely to change given the different context e.g. "*what does a plant need to grow*?" There were 13 questions in total that were only excluded from the interview schedule for the second context (presentation order of the contexts were counterbalanced). Table 7.3 below highlights these single-use questions (shaded in grey) and Table 7.4 provides more details with regards to how these single-use questions related to each biological construct. Note that in the case for Ecology, all elements were context-specific. This is also true for Evolution aside from one generic element. In contrast virtually none of the inheritance elements were context-specific.

Table 7.3. A list of the generic core knowledge elements under each biological construct

Biodiversity	Inheritance	Evolution
Q4_G_B26	Q10_G_I9	Q22_G_Ev44
Q4_G_des	Q10_G_I12	
Q7_G_B27	Q14_G_I1	
Q7_G_des	Q15_G_I2	
Q21_G_B14	Q17_G_I4	
Q22_G_B25	Q18_G_I7	
	Q18_G_I11	
	Q23_G_I5	
	Q24_G_I8	

Q24_G_I10
Q25_G_I13

After the interview was conducted, the digital recorder was stopped and the child was asked if they had any questions. The child was then debriefed, by reminding them about the aims of the study they were told about at the start of the session, thanked for their participation, and rewarded with their choice of stickers.

Table 7.4 Interview schedule in full with example answers from each cohort and related codes. Please note the coding is described in detail further below.

#	Interview Question	CKE #	CKE description & score for children's response
1	Do you think these animals live here? Why/Why not?	Ec33	Identify that animals live in habitats to which they are particularly suited
	COHORT 1: *Yeah because of the sun*		1
	COHORT 2: *erm yeah because they need hot places to live*		2
	COHORT 3: *yeah because like the lions they hunt the gazelles I think and like I'm not sure what these hunt (cheetahs) but I think they hunt the same thing and there's like zebras there*		3
2	Can all animals live here? Why/why not?	B18	Recognise different animals are found in different environments - focusing on the range of organisms supported by a habitat
		Ec30	Recognise different animals are found in different environments - focusing on the interdependence between habitat and organisms
	COHORT 1: …(shrugs)		(B18) 0 ; (Ec30) 0
	COHORT 2: *No because like polar bears they will be in the cold country because its really hot*		(B18) 2 ; (Ec30) 1
			(B18) 2 ; (Ec30) 1

	COHORT 3: *No because say there was like a polar bear it would just get too hot and die*		
3	Could this animal live somewhere else? What kind of place could it live in? Probe (go through all animals) Can they live somewhere icy like... Why/why not?	B16	Identify similarities and differences between different environments and the effect this has on the animals that live there, focusing on the range of organisms supported by a habitat
	COHORT 1: *yeah, like in deserts.*		1
	COHORT 2: *They can live in other hot places but in cold places they'll get really cold..*		2
	COHORT 3: *They could live in a zoo but only in special conditions...they can only really live here*		3
4	**What does an animal need to live? Does it get it here? Can it go somewhere else to get it? Probe**	B26 4G_des	Describe how different habitats provide for the basic needs of different kinds of animals and how they depend on each other focusing on the range of organisms supported by a habitat
	COHORT 1: water, plants, food, the sun [do these animals get that here?] yes [Can it go somewhere else to get it?] yeah.		(B26) 2 ; (4G-des) 3
	COHORT 2: *Food, water, shelter and protection.* [do these animals get that here?] *yes* [Can it go		(B26) 3 ; (4G-des) 4

	somewhere else to get it?] *no.*		
	COHORT 3: *water, food, shelter, and like...yeah that's it* [ok do the animals get that here?] *yeah but sometimes there might be very little water* [could they go somewhere else to get those things?] *they could but like if they ate meat...they might not have the certain animals to hunt...*		(B26) 3 ; (4G-des) 3
5	Why can't this [lion] live in cold weather? But a polar bear can? Why? How?	Ev46	Consider how some animals are adapted to the extreme
	COHORT 1: *because they have loads of fur to keep them warm.*		2
	COHORT 2: *maybe because they're blubber's not that thick but polar bears and penguins are* [could a polar live here?] *no it's too hot, they'd be like sweating and stuff....*		3
	COHORT 3: *because its adapted to the hot weather so like it evolved into what it is today to survive the condition.* [So what does evolved mean? or adapted?] *like evolved mean like some birds they were like,...like*		5

	some things died out cause they didn't evolve like dodo birds		
6	Can all plants live here?	Ec29	Recognise different plants are found in different environments - focusing on the interdependence between habitat and organisms
		Ec32	Identify that plants live in habitats to which they are particularly suited
		B17	Identify similarities and differences between different environments and the effect this has on the plants that live there, focusing on the range of organisms supported by a habitat
	COHORT 1: *no....because some plants will find it too dry.*		(Ec29) 0; (Ec32) 2; (B17) 0
	COHORT 2: *no because some plants need water to live*		(Ec29) 0, (Ec32) 3; (B17) 0
	COHORT 3: *no....because you need different weather for growing different things like bananas you need hot weather and like apples you need quite wet and sometimes dry weather.*		(Ec29) 2, (Ec32) 3; (B17) 1
7	What does a plant need to grow? Does it get it here?	B27 7G_des	Describe how different habitats provide for the basic needs of different kinds of plants [and how they depend on each other]
	COHORT 1: *water, food, soil* [can these plants get that		(B27) 2, (7G-des) 3

	here?] *no because they can't get soil here, but they have sand...*		
	COHORT 2: *water, sun, and soil.* [do these plants get that here?] *yeah but they don't get that much water.*		(B27) 1; (7G-des) 3
	COHORT 3: *it needs water, sunlight...and manure I think. there's quite a few animals so like...they'll be enough manure and also there's a lot of sunlight but in areas where there isn't water there'll be less plants*		(B27) 4; (7G-des) 3
8	Can they grow somewhere else like…? (Somewhere icy/cold/dry)? Why/why not?	B19	Recognise different plants are found in different environments - focusing on the range of organisms supported by a habitat
	COHORT 1: *Yes like they can grow here* (UK) [could they grow somewhere like the North pole?) *no because it's too cold*		2
	COHORT 2: *.No cause they grow in really hot places.*		2
	COHORT 3: *No because there wouldn't have any sunlight or there wouldn't be any other animals around.*		3
9	Why can't this tree grow in the cold weather, but a	Ev45	Consider how some plants and animals are adapted to the

	Christmas tree can? Why/how?		extreme
		Ev40	Recognising how plants are suited to the environment in which they live
	COHORT 1: *Because some plants like a cactus they don't need rain and like the Christmas trees, they don't need rain and they don't need sun.*		(Ev45) 1; (Ev40) 1
	COHORT 2: *because it will be too cold and it might freeze but they (xmas trees) have special roots*		(Ev45) 3; (Ev40) 3
	COHORT 3: *Because that one [points to tree] its a different type of tree than the xmas tree and it doesn't have many leaves to produce food so it makes less food*		(Ev45) 4; (Ev40) 3
10	**How does a plant change as it grows?**	I12	Describe the main stages of the plant life cycle- noting inherited traits do not change over time.
		I9	Recognise the process of growth in plants
	COHORT 1: *it gets bigger and they get small twigs*		(I12) 1; (I9) 1
	COHORT 2: *It starts to go bigger and more leaves and like the trunk of the trees gets really heavy.*		(I12) 1; (I9) 1
	COHORT 3: *first it starts off as a seed sucking up the*		(I12) 2; (I9) 3

	water and things like that, then like a seed will have a little thing come out of it to start forming its roots, then like it will have grown and it will start sucking up water to form a little sprout then it will have a leaf or something and it will start sucking up the daylight as well and then it will go eventually into its fully grown		
11	Do all plants look the same? Why/why not?	I6	Describe how offspring of plants resemble parent plants in features
		Ev42	Recognise why plants produce offspring of the same kind and link this back to inheritance
		B22	Recognise all plants show variation within the same species
		B21	Recognise all plants show variation among different species
	COHORT 1: *No because some plants will be like in different conditions so you wouldn't know which pant will need to go in a shady area or in a sunny area.*		(I6) 0; (Ev42) 0; (B22) 2; (B21) 2
	COHORT 2: *no because they are...because some plants need specific places to live than others because different cacti need really hot places, like if you put it in the garden in winter it will like get frozen up but in*		(I6) 0; (Ev42) 0; (B22) 2; (B21) 3

	the desert there's no rain COHORT 3: *No because like every tree is kinds unique like if you look for trees around you, you will never find the same kind of tree really cause they're all grown in different ways* [ok so why do they look different?] *like some plants grow on the things around them, so say like this bookcase if you grew another one outside like it won't have the shape of it*		(I6) 0; (Ev42) 1; (B22) 3; (B21) 3
12	Are all animals the same? Do they all look the same?	B23	Recognise all animals show variation among different species
	COHORT 1:*....yeah* [exactly the same?] *yeah because they're all black and white (zebras)* [what about lions?] *no because the male lions have that thing on their face*		1
	COHORT 2: *....they look the same but they're not exactly the same.* [why?] *because some zebras could be like braver and hunters and like can run faster.*		2
	COHORT 3: *....no cause they might have little differences like the zebras when you sometimes see*		3

	them they'll have like different patterns on them, they'll still be black and white but they'll have like different orders on them... [ok so what makes them have those differences?] *like the patterns on them...when they're born like they have them*		
13	Why doesn't it ever look like another kind of animal? Why does it always look like the same kind of animal?	I3 Ev43	Give reasons why living things produce offspring of the same kind Recognise why animals produce offspring of the same kind and link this back to inheritance
	COHORT 1: because other animals are not just plain black and white		(I3) 1; (Ev43) 0
	COHORT 2: *because fish has like special fins to help them live in the sea but like birds they don't have fins because they don't live in the sea.*		(I3) 3; (Ev43) 1
	COHORT 3: *because it's that species of animal*		(I3) 3; (Ev43) 3
14	**Can a __ (select specific animal) give birth to another ____ like _?**	I1	Recognise that species run true
	COHORT 1: *no because er....I'm not sure*		0
	COHORT 2: *no because it's not its breed.*		3

	COHORT 3: *no because like it couldn't give birth because its not the same species even if they're like the same family like toads and frogs, they couldn't give birth to each other because they're different types.*		4
15	**How do animals come to have babies?**	I2	Recognise the process of reproduction as the mechanism behind inheritance
	COHORT 1: er...I don't know....I think they lay eggs?		1
	COHORT 2: *erm...I think its like humans do, that they have in inside and then they just like get out...like humans do.*		2
	COHORT 3: *like they start mating then they do attempt to show off and then like...and then....I don't know.*		2
16	Does a baby ___ (select same animal as above) always look like the mum and dad ___? Why/why not?	B24	Recognise all animals show variation within the same species
	COHORT 1: erm yeah [exactly?] no it's smaller [ok would this frog grow up to look like its parents?]		1

	yeah...			
	COHORT 2: *it will look like the mum and dad but not as much. Not really as its mum and dad because its still young and it needs to grow..* [when it grows up will it look like its mum and dad?] *yes* [will it look exactly like its mum and dad?] *not exactly* [ok, but why is that?] *because erm you get the same look because it's your relative but you wont look exactly like them...its like how humans . Like my sister I won't grow up to be like here because we have to all look different.*		2	
	COHORT 3: *no* [why is that?] *cause like it might have a different pattern on it, like no animals have the same pattern* [ok, will it look a little bit like its mum and dad or not at all?] *yeah like the basic features like tails, heads.... but as it grows up it will look more like them, like birds when they hatch they look nothing like it cause they have no hairs or anything*		2	
17	**If this cheetah had a baby, will this baby look exactly**	I4		Describe how offspring of animals resemble their parents in

	like its mum and dad? Probe. Will it look more like its own mum and dad from its family, a different mum and dad from a different family, or will they all look the same?		many features
	COHORT 1: *it gets bigger, it grows more fur*		1
	COHORT 2: *he would know how to hunt for his own food and when he grows he'll get like more hair and faster tails...*		2
	COHORT 3: *it would have a different pattern but it still look the same.*		2
18	**How will this baby cheetah change as it grows up?**	I7	Recognise the process of growth in animals
		I11	Describe the main stages of the animal life cycle noting inherited traits do not change over time
	COHORT 1: *it will get bigger and it will be cleverer*		(I7) 2; (I11) 1
	COHORT 2: *well erm in how like all cheetahs looks it will look like that but it won't look exactly like the parents. .*		(I7) 2; (I11) 3
	COHORT 3: *it will get bigger and learn skills like hunting and also it would be like more aware of its*		(I7) 2; (I11) 2

	surroundings.		
19	How did the savannah/pond come to be like this?	Ev39	Identify the similarities and differences of local environments and how these affect the kinds of animals that live there [thereby shaping that environment]
	COHORT 1: *erm I don't know...*		0
	COHORT 2: *Like it was a dry landscape*		1
	COHORT 3*: Maybe because it's near the equator so it's hot*		3
20	**Do you think all [zebras] look exactly the same as each other? So why do we call them all [zebras]?**	B15	Grouping living things according to their observable similarities and differences
	COHORT 1: *because they're all zebras!*		1
	COHORT 2: because they're all the same type of animal.		3
	COHORT 3: *since they're all the same species.*		4
21	**What makes a human human?**	B14	Recognise the similarities between themselves and others

	COHORT 1: the...hand and legs and stuff [but a monkey has those...how are we different?] erm...we can like talk and stuff		1
	COHORT 2: *that a human because they're like people and not animals [*why are we different?] *because they're people...and like we don't have to hunt for our food and that we don' look all the same and...we don't need to live in a specific area.*		2
	COHORT 3: *we're they're only animals that have different languages and have like been to the moon and erm...and that invented really clever things.*		2
22	**Do people look different to each other or do they look the same?? What makes people look different/the same?**	B25 Ev44	Recognise all humans show variation within a species Recognise why humans produce offspring of the same kind and link this back to inheritance
	COHORT 1: *People look different because like....it's just like their personality*		(B25) 2; (Ev44) 0
	COHORT 2: *different* [what makes humans look different then?] *the like the countries they come*		(B25) 3; (Ev44) 3

	from.... COHORT 3: *No, even if they're twins they'll still be some differences..* [what makes humans look different then?] *like they might think in different ways or.....erm...it can like they might think what they're wearing to look a bit different*.		(B25) 2; (Ev44) 2
23	**Do people ever look like their mum and dad? Do people ever look exactly the same as their mum or dad? Why/why not? How?**	I5	Describe how offspring of humans resemble their parents in many features
	COHORT 1: just a little bit because ....I'm not sure		0
	COHORT 2: *as they grow up they might but they don't look exactly the same* [so why do we look a little bit the same as our mum and dad?] *because like we have their blood*		3
	COHORT 3: *Sometimes like you can have the features of the mum and dada like you could have your mums nose and your dads chin* [and do you know why that is?] *I don't know..*		2

24	**How does a baby change as it grows?**	I10	Describe the main stages of the human life cycle, noting inherited traits do not change over time
		I8	Recognise the process of growth in humans
	COHORT 1: *It gets bigger and it can do more stuff*		(I10) 0; (I8) 1
	COHORT 2: *It starts to grow older and gets more hairs and its body starts to change.*		(I10) 1; (I8) 1
	COHORT 3: *it learns to talk and walk and do things like erm...do its shoes up by itself and....get toilet trained and it will eventually get bigger.*		(I10) 2; (I8) 2
25	**Do animals and people grow in the same way? What is the same/different? (commonalities)**	I13	Recognise and compare the main external features of the bodies of humans and other animals
	COHORT 1: *yeah* [what's the same then?] *that they both grow [*and what's different?] *animals sometimes come from eggs but humans don't*		3
	COHORT 2: *Yes. That they like, their fur's very little and they start to do like the little things like swim and walk and they like get older and they know to like make their own food* [is there anything different?]		2

	they don't grow the same, humans have arms and they don't.			
	COHORT 3: *No like with birds, humans don't come out of eggs. But like the parents, some mammals just leave their young. It's kind of the same cause you like look after them.*		4	
26	Would it make a difference to the zebra if the lion wasn't there anymore? Vice versa. Probe.	B20	Understand the process of a feeding-chain.	
	COHORT 1: *erm yeah because the lions might eat them so they don't have to run away that much, so the zebras won' get dead and they'll be more zebras living.*		3	
	COHORT 2: *Yeah, because less lions will be hunting them down so they'll be more zebras and their breed will get bigger.*		4	
	COHORT 3: *Yeah because then it will be overpopulated with zebras and then people would have to kill them but they might kill too many so like it*		5	

	keeps the amount of zebras under control.		
27	Gazelles eat lots of grass, if there were lots of zebras around in the savannah, would it make a difference to the gazelles? Probe.	Ec31	Recognise all living things are interdependent interacting with each other and their environment (Q.29)
	COHORT 1*: yeah because there's no grass for them (zebras) to eat and they'll get hungry and they have to eat something else* [like what?] *leaves.*		3
	COHORT 2: *Yep because they'll be taking lots of its food so they won't have enough food and they'll have to like go somewhere else to find food.*		3
	COHORT 3: *No cause ...but if they ate grass then none of the grass left for them but like it might be loads of grass here but none here so the zebras will starve to death and die out.*		4
28	You see this cheetah? Is it more like the lion, more like the zebra or more like the gazelle? Why/why not? How is it similar? How is it different?	B28	Identify a number of things that can be grouped as producers, consumers, predator, prey, herbivores, carnivores etc.

	COHORT 1: *more like a lion because they're both cats*		2
	COHORT 2: *more like the lion because it doesn't eat plants it eats other animals.*		3
	COHORT 3*: like the lion because it hunts and it doesn't eat grass or anything.*		3
29	If lions eat zebras, how come there are still lots of zebras around in the savannah?	Ec31	Recognise all living things are interdependent interacting with each other and their environment (Q.27)
	COHORT 1: *because like they've got…like frogs, they can have at least 10 babies*		3
	COHORT 2: *from the zebras that still live, they might give birth to more zebras.*		3
	COHORT 3: *like they (zebras) like they give birth to young so then they can grow and look after them to grow up? Zebras give birth to it and then they grow up and then the old ones would get hunted...*		3
30	Lions eat zebras, but zebras are good at hiding so that the lions can't catch them. The zebras that have lots of stripes are good at hiding. Which zebra is the lion	Ev47 Ev41 Ev38	Recognise natural selection as the process of evolution (Q.34) Recognising how animals are suited to the environment in which they live

	likely to eat first? Why? Probe.		Identify similarities and differences between local environment and how these have an effect on the animals that live there
	COHORT 1: *The one with less stripes because they're more easy to see*.*[*right so what will happen to the numbers of those types of zebras?] *erm they will decrease* [and what about the numbers of the zebras with lots of stripes?] *they will increase.*		(Ev47) 1; (Ev38) 0; (Ev38) 0
	COHORT 2: the zebras with a little bit of stripes because its easier for them to see them and the other ones are more camouflaged [right so what's going to happen to the numbers of fish with small fins?] they'll be less of them living		(Ev47) 3; (Ev38) 0; (Ev38) 0
	COHORT 3: I think these ones because they're closest to them, or it could be that one cause it's far away from the pack so it's an easy target.		(Ev47) 2; (Ev38) 0; (Ev38) 0
31	Suppose it didn't rain for a long time, and the lake dried up. What would happen to all the animals? Probe.	Ec36	Identify the similarities and differences of different environments and how these affect the kinds of animals that live there, focusing on the interdependence between various habitats and organisms

	COHORT 1: *they'll have to go to another place to find water* [where would they go?] *I'm not sure [*ok and what would they eat somewhere else?] *grass...leaves..*		2
	COHORT 2: *they wouldn't get any water and they'll have to like...some animals they will just like sit there for rain to come and some would explore more areas to see if there's any more water? If* there's no water they will like die of thirst.		3
	COHORT 3: *they would like all erm they'll get really thirsty and they'll just die of thirst. If they was near the ocean they could get water but if not they'd just die*		3
32	What would they eat? Where would they go? Why?	Ec35	Describe how different habitats provide for the basic needs of different kinds of animals and plants, and how they depend on each other, focusing on the interdependence between habitat and organisms
	COHORT 1: *Nothing because they don't really need water*		0

	COHORT 2: *They won't have water so they will die.*		3
	COHORT 3: *It depends where they are. If they can travel to find water then it's ok but if they're in the middle of nowhere…they die.*		4
33	Suppose this lake dried up? Would this plant be able to live here? What would happen to it? Why?	Ec37	Identify similarities and differences between local environment and how these have an effect on the plants that live there focusing on the interdependence between habitat and organisms
		Ec34	Describe how diff habitats provide basic needs for plants.
	COHORT 1: *no . There's water and stuff*		(Ec37) 0; (Ec34) 1
	COHORT 2: *erm... They won't have water so they will die.*		(Ec37) 1; (Ec34) 1
	COHORT 3: *They'd all die out and lose all their leaves and like these would starve to death (zebras) and then the lions would have no food and they'd die.*		(Ec37) 3; (Ec34) 3
34	Suppose this lake dried up? Would it be better for this (choose specific animal) to have no babies, 1 baby or more babies? Why/why not? Probe.	Ev47	Recognise natural selection as the process of evolution (Q.30)
	COHORT 1: *No babies because it needs to drink and*		3

then the babies won't have any water	
COHORT 2: *erm...none because it won't have that many stuff to eat and drink.*	3
COHORT 3: *none because....actually a lot because it will eventually grow up, it wouldn't grow up very well.* [where would they get food then?] *I don't know about that...maybe lots because then can help to find food?*	4

Figure 7.2. Example of the savannah contextual scene

Figure 7.3. Example of the lake contextual scene

7.4.9.1 Development of coding scheme

A rating structure formulated by Williams and Smith (2006) was used as a basis for the development of a coding scheme, which was centred on the same core knowledge structures as the interview schedule. In their study, a content analysis was done after transcribing interviews, and key themes were taken as markers of conceptual change which were given scores. Scores given were from 1-7 where a higher score meant more conceptual knowledge. The present study therefore coded interview answers on the basis of how much core knowledge a child had about particular biological concepts; the coding generated an ordinal score and such scores were derived for each core knowledge structure. These were then totalled and an average score was obtained for each of the four biological constructs for both context-specific (asked in both interviews) and generic (asked only in the first interview) questions. Note that the core knowledge structures for generic and context-specific questions were treated separately in subsequent analyses following reliability checks for internal consistency of the scores for each biological construct, described in the next chapter.

A score of zero referred to no knowledge or irrelevance to the question. A score of one was simply a statement without any further elaboration, even after probing. A score of two was given to answers that displayed teleological, essentialist, or anthropomorphic beliefs. As such these answers could be correct, but the basis of understanding was not scientific in nature. A score of three is where answers tended to show the start of scientific understanding over social understanding, and scientific terminology begins to be used. A score of four is where this knowledge becomes more generalised in that answers are not context-specific and the child displays more consistency and confidence. Scientific knowledge is displayed, even if the mechanisms/causes/processes are unknown. A score of five is where this knowledge generally becomes more sophisticated and consistent but perhaps displaying some lack of knowledge in the mechanisms/causes/processes involved. Finally a score of six was for expert-level responses, which is a score that a primary school child is highly unlikely to obtain (Table 7.5).

One point to note is that biological constructs are to some extent interconnected and there were some core knowledge elements which related to more than one overarching construct, for example: "*Recognise all living things show variation, are interdependent, interacting with each other and their environment"* is related to biodiversity *and* ecology, and in such cases, the emphasis of the element was specified (e.g. *variation* for biodiversity versus *interdependence* for ecology) and the most direct element was scored alongside a separate score for other related elements.

Once the coding scheme had been developed, the interview answers from Pilot Study 2 were retrospectively coded to see if the system worked well enough to pursue. It was found that children's responses from the second Pilot were easily classified into one of the six codes and answers from the pilot were used as examples for future coding. Given that the sample size for Pilot Study 2 was so small (N=6), reliability checks could not be made until data from the main longitudinal experiment had been collected. These reliability checks will be described in Chapter 8.

Table 7.5. Coding scheme used to code children's responses to questions in the biological interview by cohort (C), question (Q), and core knowledge element

Score	Criteria	descriptions	examples
0	Do they address the question at all?	Don't know; silence; un-codeable; responses that indicate no knowledge or are irrelevant; no idea about entire question, no reasoning; no knowledge about mechanisms. **Note this does not refer to answers where knowledge is displayed but reasoning behind it is not.*	*…(shrugs) C1, Q2, B18* *"Hyenas are a little bit cheeky" C1, Q2, Ec30* *"I don't know" C1,C2,C3, Q14, I1* *"Not sure" C1,C2,C3, Q34, Ev47*
1	Do they address the question by making a simple assertion? Maybe a description only	Fact; tautological; non-exploratory simple assertions (possibly inaccurate); no list knowledge (e.g. quantifying the things a plant needs to grow); reasoning behind their answer is limited or not offered or based on media/observation or simple deduction based on the picture.	*"yeah, like in the deserts" C2, Q3, B16* *"They will die." C3, Q32, Ec35* *"It gets bigger and they get small twigs" C1, Q10, I12* *"No [why not? because I saw it on TV" C2, Q11, Ev42*

2	Do they address the question by using some kind of social reasoning	Inaccurate or little factual knowledge but an attempt at an explanation. Here they try to offer an answer/explanation but without coherence, ideas are context-specific and not generalised. Creative reasoning/explanations used often based on material they have in front of them (often not logical or accurate). Social reasoning such as essentialism and anthropomorphism may be prevalent (these again may have no logic behind them) and explanations are not biological in nature at all. List knowledge (e.g. quantifying what a plant needs to grow) may include <2 correct answers. Theological explanations/reasoning offered. None or very few correct 'life stages' recognised and none are explained.	*"people look different because it's just their personality" C1, Q22, B25* *"Lions will be confused because of the stripes" C2, Q29, Ec31* *"If humans don't look the same that means animals won't either" C2, Q17, I4* *"It's used to it" C1, Q5, Ev46*
3	Do they address the question by using slightly more biological reasoning or an attempt at explaining a biological mechanism of	Better consistency in explanations but ideas are still not altogether generalised, or always accurate. Perhaps still some evidence of creative reasoning but the start of more coherent logical explanations. Knowledge about mechanisms is still very little or unknown but an attempt at an explanation is made,	*"Yeah, because less lions will be hunting them down so they'll be more zebras and their breed will get bigger." C2, Q26,, B20* *"because there's no grass for them to eat and*

	sorts?	usually a combination of social and biological reasoning, but social reasoning is more prevalent (more logic to this reasoning than above). Scientific vocabulary rarely used (<1term) and list knowledge (e.g. about the number of things a plant needs to grow) <3 correct items. Theological explanations occasionally offered for complex mechanisms. Only some (<3) correct 'life stages' recognised but these are not explained. Children answers are fixated upon surface traits.	*they'll get hungry and have to eat something else" C3, Q27, Ec31* *"animals sometimes come from eggs but humans don't" C1, Q25, I13* *"It's near the equator" C3, Q19, Ev39*
4	Do they answer the question using more biological reasoning and is this more generalised?	Reasoning/explanations are much more generalised across ideas and there is increasing consistency and confidence in answers. Some scientific terms are starting to be used (<2) possibly incorrectly and they may be misunderstood. The use of more biological reasoning (some inaccuracies or inappropriate) *over* social reasoning. Mechanisms where appropriate are inaccurate or unknown but seemingly more logical in nature (explanations at this point may be more descriptive in nature given their limited knowledge). List knowledge (e.g. number of things a	*"Animals that eat grass hibernate I think…" C2, Q4, B26* *"They'd all die out and lose all their leaves and like these would starve to death (zebras) because they need plants and then the lions would have no food if the zebras die and they'd die too. It's like a chain."* C3, Q33, Ec34 *"Because if I look whole,*

		plan needs to grow) of <4 correct answers. No reciprocity in thought, knowledge is still piecemeal. Majority of (>4) life stages described but very few explained.	*like all of it like mum, I'd be just like her. If I looked all of it like my dad, I'd look like him. So if it's half and half, I'm him and her." C3, Q23, I5* "*I think these ones (zebras) because they're closest to them (lions), or it could be that one cause its far away from the pack so it's an easy target."* C3, Q30, Ev47
5	Do they answer demonstrating a more generalised and holistic understanding?	The start of more biological *causal* explanations, which are consistent across contexts. Some use of scientific/biological words (> 3 words), but still possibly misunderstood. Mechanisms, where appropriate, may be more logical than above but possibly still inaccurate/inappropriate. More coherent and/or explicit understanding is demonstrated (i.e. responses do not seem as though they are merely guesses, and causal explanations are offered as a basis for biological reasoning). There is	" It would make a big difference because then there would be a large overpopulation of zebras, so then there'd be no grass left, so then the zebras will all die out! So you need a little of predetors." C3, Q6, B20. *"It's a good place to keep her babies safe from any predators on land, cause if the frog had her babies on land there might be*

	reciprocity in thought and systems are understood in more than just one-way relationships. List knowledge is good <6 correct. Some details in mechanism knowledge may be unknown/misunderstood but generally the broader ideas are grasped. Nearly all ‘life stages’ are recognised and explained.	*some predators on land who might want to eat it and they have some soft jelly around it, cause I felt one in the lake before and umm so they need water to keep their surface slippery or otherwise they’ll dry up and die.” C3, Q1, Ec33.* *“Their cells join together and they can erm make a baby so it basically makes a mixture of them.” C3, Q15, I2* *“because it’s adapted to the hot weather so like it evolved into what it is today to survive the condition.* [So what does evolved mean? or adapted?] *like evolved mean like some birds they were …some things died out cause they didn’t evolve like dodo birds” C3, Q5, Ev46*

6	Expert level answers. Very little incorrect.	Better grasp of interactions/interdependency/ reciprocity. Genetic reasoning or correct biological explanations are offered where appropriate; dimensionality of time may still be an issue. Very little that is incorrect.	*No children displayed this code in the first pilot study so there are no examples.

Note that the scale above is ordinal and not interval (which would be required for the planned longitudinal data analysis. However, there is substantial precedent for treating data from scales such as these as interval. For example science-related studies by Williams and Smith (2010), Myant and Williams, (2005), Howe et al., (2005), Howe et al (2012) and Phillips and Tolmie (2007) have used ordinal scores, which are later transfored into quasi-interval scores in order to conduct data analysis.

Interval scores are preferable because they allow the use of more powerful and systematic analysis of patterns of data using parametric techniques. Treating ordinal scores as interval scores is not a violation of statistical principles despite this often being a concern, however as Hair et al (2009) describe, parametric techniques essentially require estimates of central tendency to be normally distributed rather than the actual data points, thus as long as there are no obvious distortions in the data (e.g. floor or ceiling effects) these techniques tend to be largely robust.

Nonetheless, something to note is that in the present study, this does not mean that there is not some tension between the detail of children's responses and the reduction of these to interval indices, however the objectives of the research in terms of looking at correspondences between areas of understanding necessitated an emphasis on interval scores given the planned data analysis, and more qualitative aspects of the data can be explored elsewhere. The ordinal raw scores capture conceptual progression and there is in fact some tendency towards normal distribution, with relatively few responses at the extremes of the scale, although it is necessary to consider what increases in score imply a shift from and to in terms of content, when addressing their interpretation.

The examples shown in the interview schedule (Table 7.4) of the average responses from each cohort for each core knowledge element tells us some useful things. For example, the key areas missing from children's knowledge seemed to be ideas around causal process and mechanisms behind biological phenomena, and reasons why things were as they were. Often young children would wildly guess the answers to questions addressing mechanistic understanding. It was only toward Year 5 that some children were able to offer logical and at times correct answers. This suggests that children's early knowledge may not be theoretical, but perhaps more perceptual in nature, although a larger sample size would be needed to confirm this.

In fact there might be a change in the type of causal reasoning children used with age. Many of the explanations behind children's ideas in Year 1 were attributable to more psychological or teleological reasoning, rather than mechanistic or biological causal explanations offered by children in Year 5. For example, when children were asked what the effect would be if a lion no longer lived in the savannah, a boy aged 6 replied: "*zebras would be happier because lions used to scare them*." In contrast, a boy aged 10 replied *"if there was no lions whatsoever then the zebras could populate a little bit too much and that's when* *the lion needs to come back"*, thus demonstrating key differences in their reasoning. A key turning point for social to biological causal reasoning seemed to be around Year 3 where children preferred biological explanations *over* social ones. This is something the coding scheme is also able to detect effectively. However despite this, some social reasoning was still offered even among the older children in Year 5. It may be that patterns of non-biological reasoning are based around observation and discourse within a child's own personal experience (cf. Frusal, 2015). This is further supported by the fact that it was found media and prior experience seemed to heavily influence children's knowledge, particularly when they lacked biological understanding. These ideas will need to be investigated in more detail in the main experiment, however for now, the findings indicate sensitivity on part of the coding scheme and interview schedule.

7.4.10 Overall procedure

Upon entering the testing room, children were informed about the study and given the opportunity to ask questions. Children were told that the experimenter was interested in how children across primary school were learning science, and that today they were going to be playing some games that were related to science learning. They were assured that the experiment was in no way compulsory, was not related to their academic achievement and that results would remain confidential. They were then asked for their verbal consent and if obtained, details about the child's name (each child was later assigned an anonymous identifying code, but the child's name was needed to be able to follow them up the following year and to ensure parental consent), age, and class. The child was then given instructions for each subsequent cognitive tests given described earlier in section 7.4.1 and were not told their results on any of the tests, but were given praise throughout such as "*well done, you're doing really well*." The child had the opportunity to ask questions throughout the testing period and was debriefed at the end of the session.

The order of the measures for general cognitive abilities was organised in such a way so as to ensure no boredom or fatigue among children. The order of the tasks was always the same: WCST part one, Stroop task set A, digit recall, block recall, backwards digit recall, BPVS, NKT, Stroop task set B, and WCST part two (stoop task was set A and set B were counterbalanced). Average testing time was approximately 30-40 minutes for each child for the first session on the general cognitive ability measures. The second session for the biological knowledge task was conducted approximately three weeks later and also took around 30 minutes per child.

All testing was conducted in a quiet meeting room within the school, with the researcher and the child alone. A table with the materials was set up beforehand and the child sat to the left of the experimenter in front of the table.

7.5. Phase Two

7.5.1 Participants

Participants that were recruited in Phase One of the experiment were followed up in Phase Two from the same three state schools in London. As parental consent had already been obtained for the entire longitudinal study, only verbal consent of the child was obtained prior to testing. Given that London schools typically have a high turnover, attrition was expected. A total of nine children had left the study, leaving the total number of children at Phase Two at 129 making up 6.5% of the original sample. Details of attrition levels are shown in Table 7.6 below:

Table 7.6. Rates of attrition in each Year group across the three schools recruited in this study

		School			
		OG	GT	BP	Total
Year at school	Year 1	1	2	2	5
	Year 3	0	1	1	2
	Year 5	0	2	0	2
	Total	1	5	3	9

This phase of the study aimed to test children from Year 1 (age 5/6 years), Year 3 (age 7/8), and Year 5 (age 9/10). There were 45 participants from school OG, 42 participants from school GT and 42 participants from school BP. Of the total sample, 38 children are in Year 1 (mean age = 57.05 months, SD = 3.53), 41 from Year 3 (mean age = 81.12 months, SD = 7.27) and 41 from Year 5 (mean age = 115.30 months, SD = 3.67). Once again the standard deviation for Cohort 2 was more widely distributed than for the other two cohorts suggesting a wider distribution of ages around the mean.

7.5.2 Materials & administration Phase Two

The materials and procedure for Phase Two session one on the general cognitive measures were exactly the same as those described above for Phase One session one. Likewise the materials and the procedure for the biological task were exactly the same as those in Phase One, except with the addition of one more task at the end.

The vocabulary subtest from the Wechsler Abbreviated Scales of Intelligence (WASI; Wechsler, 1999) test battery was included in the study as a measure of expressive language. It was felt necessary to include this task given that the biological task itself was expressive in nature. Unfortunately this test could not have been used during Phase One because it could only be administered to children aged 6 years and above. As children were older, this measure was subsequently included at the very end of testing after the biological measure. The reasons for this were so that the testing time for session one would not be extended, but also because it ensured consistency of testing sequence between Phase One and Phase Two up to this point. If there were any interaction between testing order and performance, this would therefore remain stable.

7.5.2.1 Expressive language

Children were told that they would hear the experimenter say a word and that they were to tell the experimenter the meaning of the word as best they could: "*I'm going to say a word, and all I'd like you to do is to listen carefully and tell me what you think that word means."* If children were stuck they were prompted to give an example or to use the word in a sentence (as specified by the instructions in the WASI manual). Words used in the task increased in difficulty and children's answers were coded as either zero for incorrect, one for partially correct, and two for correct. Children started the task at various stages, following the guidelines of the WASI manual. If they failed to score two for the first two trials, the task was administered backwards until a baseline could be reached (score of two for two consecutive trials). For the children aged six, or those that failed the first block of verbal trials, the first four trials consisted of showing children a picture of a shell, shovel, fish, and map which the child was asked to label correctly. These were scored in the same way as the other trials. If children failed to answer correctly five consecutive times, the task was stopped, and their scores totalled up. The final overall score was used as a measure of expressive language.

7.5.3 Overall procedure for Phase Two

The overall procedure for Phase Two was exactly the same as the procedure described in Phase One (see section 7.4.2) with the addition of the expressive language task in session two, described above. Testing took place in the same meeting rooms in each school, and the materials were set up in the same way as in Phase One, with the child always sitting to the left of the experimenter. Testing time was 30-40 minutes for each session.

Children were fully debriefed at the end of the final testing session and were told the aim of the study was to try to understand how children in primary school were learning science because some children find science a difficult subject to learn. They were also told how learning changed from Reception all the way through to Year 6, which is why children across a number of year groups took part. Children were made aware that the tasks they participated were all linked to science learning in some way and were then given the opportunity to ask any specific questions, which the experimenter answered in as much detail as possible. Children were once again reminded that their data would remain confidential and anonymous, and they were then thanked for their participation and given a selection of stickers as a token of gratitude.

7.6. Teacher data

It was thought that obtaining some background information about the kinds of concepts children are exposed to in class across the three schools, and the types of teaching strategies used would be valuable information to have. A questionnaire was developed which aimed to collect the said information as well as the education level of the teachers, their professional training, and confidence levels of teaching science at primary school. These questionnaires were administered to teachers across all year groups at all three schools in the study. Teachers were told their participation was absolutely voluntary, and that information they supplied would remain anonymous, strictly confidential, and if they later chose to withdraw their data, they were able to do so with immediate effect. A copy of the teacher questionnaire can be found in the appendix (A.4).

Unfortunately the return rate of completed questionnaires was very low, with only 13 questionnaires having been returned (see Table 7.7). At least one questionnaire was returned for each academic year group apart from Reception where no teachers returned their questionnaires. For these reasons data were explored qualitatively in the appendix (A.5) and interpreted with caution.

Table 7.7. Number of teacher questionnaires returned by each school alongside the year group each teacher taught

School	Number of Teachers	Number of returned questionnaires
OG	14	0
GT	19	7: 2 x Year 1; 1 x Year 2; 1 x Year 3; 2 x Year 4; 1 x Year 6
BP	32	6: 1 x Year 1; 1 x Year 3; 1 x Year 4; 1 x Year 5; 2 x Year 6

CHAPTER 8: RESULTS TIME 1

8.1 Overview and analysis strategy

The data created by the triple cohort longitudinal design made it possible to examine general patterns of emergence and change in biological constructs, but also to derive statistical models of the connections over time between different areas of understanding, and the influence of developing general cognitive abilities. Data coding took place throughout the period of data collection, with analysis starting after the completion of Phase One. Whilst data collection was on-going, analysis was generally cross-sectional, with the focus on identifying and comparing performance within each cohort, and examining associations between different variables.

This chapter outlines the analyses computed[3] after Phase One (henceforth also referred to as Time 1) of the data collection. The chapter starts by addressing the reliability of the biological interviews and the internal consistency of the biological constructs. Note that the psychometric properties of the general cognitive measures were not assessed since these were derived from standardised tests.

The exploratory descriptive analyses for general cognitive and biological measures are then presented, before assessing the significant differences in performance between cohorts, gender, order of presentation of the biological task, contextual scene, and age differences across all measures.

[3] All data were analysed using IBM Statistics version 22 and Microsoft Excel version 2013.

Associations between all variables were also explored using correlational analyses. These were conducted to examine the relationship between the demographic data and the biological performance data, which made it possible to identify both overall trajectories and sub-groups with specific developmental patterns, along with predictors of these variations. Lastly in this chapter, the data obtained from teacher questionnaires are examined.

8.2 Reliability of coding system for biological task

Development of the coding scheme is described in Chapter 7. In order to check the reliability of the coding system, two coders independently coded children's responses to the biological interview from a random sample of 12 transcribed interviews, using the scoring rubrics described previously. Four children were randomly selected from each cohort (which the coders were blind to), with equal males and females. Reliability of scoring was assessed by examining the percentage agreement scores between two raters for each core knowledge structure for each child. Note that given the nature of the interview schedule, there were a number of interview questions, and therefore core knowledge structures, that were context-specific and as a result these were only asked once. There were a total of 19 context-specific elements and 64 non-context-specific or 'generic' items; hence an overall total of 83 scores per child (see Table 8.1 for a list of generic questions).

Cohen's kappa was not applicable in this case because the six level scoring multiplied by two raters would create too many analyses, which would have required an unfeasibly large sample to populate reliably. Secondly, the large number of analyses would have created substantial possibility of chance effects; hence percentage agreement scores were used as an alternative in two ways: across elements within each participant (Table 8.2), and across participants within each element (Table 8.3). Taken together, these provided a good triangulation on how far independent scoring consistently arrived at the same results. Table 8.3 shows the percentage agreement values between raters across the total number of core knowledge structures for each randomly selected participant. The range of percentage agreement was very narrow at 13.25%, and the average was high at 72.69% indicating acceptable levels of reliability in coding.

Table 8.1. Single-use questions used in the biological interview schedule (total 19)

Question		Core Knowledge Element	
4	What does an animal need to live? Probe	B26	Describe how different habitats provide for the basic needs of different kinds of animals and how they depend on each other focusing on the range of organisms supported by a habitat
		Q4_des	List a number of correct items that animals need to survive
7	What does a plant need to grow? Probe	B27	Describe how different habitats provide for the basic needs of different kinds of plants [and how they depend on each other]
		Q7_des	List a number of correct items that plants need to survive
10	How does a plant change as it grows?	I12	Describe the main stages of the plant life cycle- noting inherited traits do not change over time
		I9	Recognise the process of growth in plants
14	Can a __*(select specific animal)* give birth to another ____ like _?	I1	Recognise that species run true

15	How do animals come to have babies?	I2	Recognise the process of reproduction as the mechanism behind inheritance
17	If this___ had a baby, will this baby look exactly like its mum and dad? Probe. Will it look more like its own mum and dad from its family, a different mum and dad from a different family, or will they all look the same?	I4	Describe how offspring of animals resemble their parents in many features
18	How will this baby ___change as it grows up?	I7	Recognise the process of growth in animals
		I11	Describe the main stages of the animal life cycle *noting inherited traits do not change over time*
20	So why do we call them all zebras/fish?	B15	Grouping living things according to their observable similarities and differences
21	What makes a human human?	B14	Recognise the similarities between themselves and others
22	Do people look different to each other or do they look the same?? What makes people look different/the same?	B25	Recognise all humans show variation within a species
		Ev44	Recognise why humans produce offspring of the same kind and link this back to inheritance
23	Do people ever look like their mum and dad? Do people ever look exactly the same as their mum or dad? Why/why not? How?	I5	Describe how offspring of humans resemble their parents in many features
24	How does a baby change as it grows?	I10	Describe the main stages of the human life cycle, *noting inherited traits do not change over time*
		I8	Recognise the process of growth in humans

25	Do animals and people grow in the same way?	I13	Recognise and compare the main external features of the bodies of humans and other animals

Table 8.2. Percentage agreement of coders from two independent raters on a sample of biological interviews for each participant

Participant	Total judgements	Number of disagreements between raters' scores	% Agreement
9	83	19	77.108
14	83	24	71.084
24	83	27	67.470
26	83	16	80.723
53	83	23	72.289
65	83	22	73.494
67	83	25	69.880
81	83	21	74.699
108	83	26	68.675
113	83	24	71.084
121	83	22	73.494
123	83	23	72.289
Total	996	272	72.691

With regards to reliability across elements within each participant, percentage agreement was very high at 86.195% on average (see Table 8.3), which suggested substantially robust consistency of scoring.

Table 8.3. Percentage agreement of raters for every biological element scored across a random sample of 12 participants

Core Knowledge Element	Total judgements	Number of disagreements	% Agreement
Q1_pon_Ec33	24	1	95.833
Q1_sav_Ec33	24	3	87.500
Q2_pon_B18	24	5	79.167
Q2_sav_B18	24	5	79.167
Q2_pon_Ec30	24	4	83.333
Q2_sav_Ec30	24	5	79.167
Q3_pon_B16	24	3	87.500
Q3_sav_B16	24	6	75.000
Q4_G_B26	24	4	83.333
Q4_G_des	24	2	91.667
Q5_pon_Ev46	24	4	83.333
Q5_sav_Ev46	24	6	75.000
Q6_pon_Ec29	24	3	87.500
Q6_sav_Ec29	24	2	91.667
Q6_pon_Ec32	24	3	87.500
Q6_sav_Ec32	24	5	79.167
Q6_pon_B17	24	4	83.333
Q6_sav_B17	24	4	83.333
Q7_G_B27	24	6	75.000
Q7_G_des	24	3	87.500
Q8_pon_B19	24	3	87.500
Q8_sav_B19	24	5	79.167
Q9_pon_Ev45	24	5	79.167
Q9_sav_Ev45	24	5	79.167
Q9_pon_Ev40	24	4	83.333
Q9_sav_Ev40	24	3	87.500
Q10_G_I9	24	6	75.000
Q10_G_I12	24	1	95.833

Q11_pon_I6	24	2	91.667
Q11_sav_I6	24	1	95.833
Q11_pon_Ev42	24	2	91.667
Q11_sav_Ev42	24	1	95.833
Q11_pon_B22	24	5	79.167
Q11_sav_B22	24	3	87.500
Q11_pon_B21	24	5	79.167
Q11_sav_B21	24	5	79.167
Q12_pon_B23	24	4	83.333
Q12_sav_B23	24	3	87.500
Q13_pon_I3	24	5	79.167
Q13_sav_I3	24	5	79.167
Q13_pon_Ev43	24	2	91.667
Q13_sav_Ev43	24	4	83.333
Q14_G_I1	24	5	79.167
Q15_G_I2	24	6	75.000
Q16_pon_B24	24	3	87.500
Q16_sav_B24	24	5	79.167
Q17_G_I4	24	6	75.000
Q18_G_17	24	5	79.167
Q18_G_I11	24	1	95.833
Q19_pon_Ev39	24	3	87.500
Q19_sav_Ev39	24	4	83.333
Q20_G_B15	24	4	83.333
Q21_G_B14	24	4	83.333
Q22_G_B25	24	2	91.667
Q22_G_EV44	24	2	91.667
Q23_G_I5	24	6	75.000
Q24_G_I8	24	5	79.167
Q24_G_I10	24	2	91.667
Q25_G_113	24	0	100.000
Q26_pon_B20	24	5	79.167

Q26_sav_B20	24	5	79.167
Q27_pon_Ec31	24	4	83.333
Q27_sav_Ec31	24	5	79.167
Q28_pon_B28	24	3	87.500
Q28_sav_B28	24	2	91.667
Q29_pon_Ec31	24	2	91.667
Q29_sav_Ec31	24	3	87.500
Q30_pon_Ev47	24	3	87.500
Q30_sav_Ev47	24	1	95.833
Q30_pon_Ev41	24	2	91.667
Q30_sav_Ev41	24	3	87.500
Q30_pon_Ev38	24	3	87.500
Q30_sav_Ev38	24	3	87.500
Q31_pon_Ec36	24	1	95.833
Q31_sav_Ec36	24	2	91.667
Q32_pon_Ec35	24	0	100.000
Q32_sav_Ec35	24	1	95.833
Q33_pon_Ec37	24	0	100.000
Q33_sav_Ec37	24	3	87.500
Q33_pon_Ec34	24	2	91.667
Q33_sav_Ec34	24	0	100.000
Q34_pon_Ev47	24	1	95.833
Q34_sav_Ev47	24	1	95.833
		Total % average	86.195

8.2.2 Internal consistency of the biological constructs

The internal consistency of the scores for each biological construct was assessed using Cronbach's alpha. This was done first by checking the consistency for each construct within context, and if this appeared to be reliable, the consistency was also checked across context. The total scores for each biological concept (biodiversity, ecology, inheritance, and evolution) were computed by totalling the scores each child obtained for each individual core knowledge structure within a given biological construct. As there were two contexts, there were some core knowledge structures for which children had no score, as questions were not always repeated for each context. This meant that an alpha could not be computed for all the core knowledge structures within a biological construct. Also, because of the way the interview was developed, questions that were perceptually context-specific were asked twice, once for each context, whereas questions that were more generic in nature and perceptually not related to either of the contexts presented to the children, were only asked once to avoid repetition. For this reason, the context-specific questions and the generic questions were treated separately in the reliability analysis. Cronbach's alpha scores were computed for each biological construct on context-specific core knowledge structures for the pond context, and then for the savannah context. The results of these analyses are in Table 8.4.

Table 8.4. Cronbach's alpha scores for context-specific core knowledge elements in each context

context-specific elements				context-specific elements			
Time 1		alpha	items	Time 2		alpha	items
Biodiversity	pond	0.773	11	Biodiversity	pond	0.816	11
	savannah	0.736	11		savannah	0.882	11
Ecology	Pond	0.691	10	Ecology	Pond	0.677	10
	Savannah	0.682	10		Savannah	0.699	10
Inheritance	Pond	0	2	Inheritance	Pond	0.103	2
	Savannah	0	2		Savannah	0.234	2
Evolution	Pond	0.611	10	Evolution	Pond	0.660	10
	Savannah	0.616	10		Savannah	0.665	10

Aside from inheritance concepts, the table highlights that these analyses generally produced high alpha scores of >0.6 and values were similar across contexts. Alphas for inheritance concepts were quite low; however, the item scoring suggested that children generally scored “0” on the context-specific elements for inheritance, of which there were only two items, and alphas for small item sets tend to be lower. The fact that children are showing a lack of knowledge on inheritance concepts early on is interesting, and will be explored further on.

Given the alpha scores for the other biological constructs were relatively high, there was a case for computing a *total score* for each biological construct based on: context-specific pond mean scores, context-specific savannah mean scores, and generic mean scores. These composite alphas for these more broad-based scores are shown below in Table 8.5:

Table 8.5. Composite alphas for each biological construct

Composite Scores	Time 1	Time 2	items
Biodiversity	0.846	0.858	27
Inheritance	0.856	0.926	15
Ecology	0.621	0.765	20
Evolution	0.608	0.781	21

The alphas obtained from the composite scores suggest high reliability in the scale used to test children’s biological knowledge, and the items used to investigate each biological construct. For these reasons all subsequent analyses were conducted using composite mean scores for all biological constructs at Time 1 and Time 2 because the different number of items for each biological construct meant the raw totals were not directly comparable.

8.3 Descriptive analyses

The four variables for biological constructs, and the seven variables for the general cognitive ability measures were analysed to observe performance across cohorts at Time 1. Data on the biological constructs and the general cognitive measures were tested for normality. The Kolmogorov-Smirnov and Shapiro-Wilk statistics are shown in Table 8.6:

Table 8.6. Results from tests of normality for biological constructs and general cognitive ability measures

Tests of Normality	Cohort	Kolmogorov-Smirnov		Shapiro-Wilk	
		Statistic	df	Statistic	df
Biodiversity	Reception	0.061	47	0.983	47
	Year 2	0.068	44	0.984	44
	Year 5	0.085	45	0.982	45
	Whole sample	0.079*	136	0.974*	136
Ecology	Reception	0.163*	47	0.937*	47
	Year 2	0.143*	44	0.928*	44
	Year 5	0.103	45	0.942*	45
	Whole sample	0.119**	136	0.905**	136
Inheritance	Reception	0.188**	47	0.897**	47
	Year 2	0.117	44	0.933*	44
	Year 5	0.196**	45	0.916*	45
	Whole sample	0.090*	136	0.914**	136
Evolution	Reception	0.120*	47	0.960	47
	Year 2	0.117	44	0.922*	44

	Year 5	0.106	45	0.963	45
	Whole sample	0.087*	136	0.961**	136
Number knowledge	Reception	0.133*	47	0.948*	47
	Year 2	0.114	44	0.959	44
	Year 5	0.169*	45	0.902**	45
	Whole sample	0.066	136	0.990*	136
BPVS	Reception	0.076	47	0.971	47
	Year 2	0.059	44	0.963	44
	Year 5	0.099	45	0.973	45
	Whole sample	0.055	136	0.991	136
Digit Recall	Reception	0.137*	47	0.973	47
	Year 2	0.209**	44	0.921*	44
	Year 5	0.084	45	0.979	45
	Whole sample	0.136**	136	0.970*	136
B/W Digit Recall	Reception	0.195*	47	0.938*	47
	Year 2	0.136*	44	0.896**	44
	Year 5	0.189**	45	0.946*	45
	Whole sample	0.154**	136	0.927**	136
Block Recall	Reception	0.125	47	0.965	47
	Year 2	0.121	44	0.963	44
	Year 5	0.111	45	0.978	45
	Whole sample	0.055	136	0.990	136
WCST	Reception	0.212**	47	0.863**	47
	Year 2	0.162*	44	0.871**	44
	Year 5	0.171*	45	0.936*	45
	Whole sample	0.205**	136	0813**	136
Stroop	Reception	0.243**	47	0.787**	47
	Year 2	0.265**	44	0.751**	44
	Year 5	0.382**	45	0.673**	45
	Whole sample	0.264**	136	0.695**	136

*p<0.05; **p<0.001

The tests of normality in Table 8.6 suggest the majority of the data were not normally distributed across cohorts as highlighted by the significant values, aside from biodiversity scores, receptive language (BPVS), and block recall scores, where the values are not significant. These significant values tend to be concentrated among the younger children, suggesting that they are primarily the result of floor effects. Also, the skewness and kurtosis values (Tables 8.7 and 8.8) suggest that there are very few extreme departures from the mean values for the majority of the test measures, and regardless, the sample size for the current study is large enough to proceed with parametric analyses.

Below is a table of means and standard deviations of children's biological performance at Time 1. It can be seen that with increasing age, children's performance improves, as would be expected. It also suggests that biodiversity is the area where children have the most knowledge, even at school entry. In comparison, inheritance is the area where children seem to have the least knowledge.

Table 8.7. Means and standard deviation statistics for performance on each biological construct by cohort

		Mean	SD	Skewness	Kurtosis
Biodiversity	Reception	1.641	0.507	0.198	-0.396
	Year 2	2.763	0.725	0.337	-0.177
	Year 5	3.804	0.745	-0.018	-0.632
	Whole sample	2.719	1.110	0.267	-0.802
Ecology	Reception	1.038	0.404	0.963	1.212
	Year 2	1.527	0.615	0.908	0.821
	Year 5	2.278	0.898	0.815	0.406
	Whole sample	1.607	0.839	1.224	1.616
Inheritance	Reception	0.609	0.432	0.506	-1.111

	Year 2	1.161	0.618	0.841	0.389
	Year 5	1.554	0.842	0.853	0.184
	Whole sample	1.1004	0.755	1.097	1.258
Evolution	Reception	0.896	0.500	0.583	0.090
	Year 2	1.643	0.796	1.258	2.831
	Year 5	2.324	0.773	0.705	0.746
	Whole sample	1.610	0.911	0.761	0.605

Table 8.8 shows means and standard deviations of children's performance on the measures of general cognitive ability at Time 1. As expected, children improved their performance on these tasks with increasing age, also.

Table 8.8. Means and standard deviation statistics for performance on each general cognitive measure by cohort

		Mean	SD	Skewness	Kurtosis
BPVS	Reception	63.230	17.533	-0.441	0.307
	Year 2	86.560	13.910	-0.710	1.299
	Year 5	121.360	16.472	-0.002	-0.741
	Whole sample	90.140	28.896	0.052	-0.424
Digit Recall	Reception	22.960	3.810	0.364	0.549
	Year 2	27.870	4.556	0.859	0.338
	Year 5	32.620	5.462	0.390	-0.101
	Whole sample	27.760	6.107	0.556	0.068
B/W digit recall	Reception	5.470	1.544	0.007	1.634
	Year 2	9.980	3.625	0.915	0.445
	Year 5	14.780	3.855	0.204	-1.050
	Whole sample	10.010	4.960	0.652	-0.529
Block recall	Reception	16.890	3.908	-0.487	1.672
	Year 2	20.780	3.437	-0.324	-0.297

	Year 5	25.220	3.971	-0.140	0.707
	Whole sample	20.890	5.097	-0.068	0.223
NKT	Reception	77.997	6.775	-0.543	-0.387
	Year 2	78.211	8.922	-0.772	1.561
	Year 5	91.599	6.230	-0.924	0.112
	Whole sample	82.570	9.740	-0.354	0.077
WCST	Reception	14.810	7.503	1.176	0.606
	Year 2	10.440	4.256	1.730	5.797
	Year 5	9.220	2.875	0.331	1.810
	Whole sample	11.580	5.784	1.891	4.217
Stroop task	Reception	2.430	2.717	1.960	4.666
	Year 2	1.390	1.919	1.443	1.268
	Year 5	0.420	0.543	.768	-0.536
	Whole sample	1.430	2.114	2.480	8.037

8.4 Significance testing

8.4.1 Gender

There were no significant differences in gender for any of the general cognitive ability measures, or the biological constructs; hence gender differences are not examined further at Time 1.

8.4.2 Order of presentation of context

Tests were carried out to see if there were any significant differences in children's performance of the biological tasks for each construct, depending on the order of presentation of the contexts. An analysis of variance (ANOVA) revealed no significant order effects for all biological constructs except for inheritance concepts: $F(126)=4.710$, $p=0.032$, two-tailed where children performed better in the second context presented (M=1.242, SD=0.757), over the first (M=0.956, SD=0.735). This was examined further by looking at cohort differences in presentation order, and these analyses revealed that there were order effects only for children in Cohort 1 (Reception): $F(41)=10.216$, $p=0.003$, two-tailed, where children generally performed better in the second context presented (M=1.161, SD=0.479) than the first (M=1.113, SD=0.677); but not for children in Cohorts 2, or 3.

8.4.3 Context

Differences in children's performance of the biological task for each context were also assessed. T-tests revealed no significant differences in children's performance for either context with the exception of inheritance, where children performed significantly better on the savannah context (M=0.343, SD=0.387) than in the pond context (M=0.277, SD=0.363): $t(136)=-2.062$, $p=0.044$ (two-tailed). When this was explored further between cohorts, it was only children in the second cohort (Year 2) displayed better performance on the savannah context for inheritance (M=0.386, SD=0.355) over the pond context (M=0.284, SD=0.312): $t(43)=-2.285$, $p=0.027$, two-tailed. The remaining two cohorts did not display any contextual differences for any of the other biological concepts.

8.4.4 Age differences in general cognitive ability measures

A 7x3 factorial ANOVA revealed a significant main effect of year group on performance on all the tests of general cognitive ability (Table 8.9). Bonferroni *post-hoc* tests revealed there were significant differences across all year groups for BPVS, digit recall, backwards digit recall, and block recall (all $p<0.001$) where children in Year 5 performed better than children in Year 2, who performed better than children in Reception. For the NKT, there was only a significant difference between Year 5 (M=91.599, SD=6.230) and Reception (M=16.894, SD=6.775), and Year 5 and Year 2 (M=78.211, SD=6.230), both at $p<0.001$. There was no significant difference in NKT between Reception and Year 2 ($p=1$) suggesting these two cohorts performed alike. The similar story emerged for WCST, where there were significant differences between Reception (M=14.809, SD=7.503) and Year 2 (M=10.444, SD=4.256), and Reception and Year 5 (M=9.222, SD=2.875) both at $p<0.001$, but not between Year 2 and Year 5 ($p=0.825$) suggesting children in the older two cohorts made a similar number of perseverative errors which were significantly fewer than children in Reception. Lastly, for the Stroop task, there was a significant difference between Reception (M=2.426, SD=2.717) and Year 2 (M=1.378, SD=1.898) at $p=0.034$, and between Reception and Year 5 (M=0.422, SD=0.543) at $p<0.001$. There was no significant difference between Year 2 and Year 5 ($p=0.065$) suggesting the children in the older cohorts made a similar number of incongruent errors which were significantly fewer than children in Reception. The results for WCST and Stroop task may indicate ceiling levels for the older two cohorts.

Table 8.9. ANOVA for effect of year group on tests of general cognitive ability

	F	*p*	partial η^2
BPVS	151.410	<0.001	0.693
Digit recall	49.743	<0.001	0.426
b/w digit recall	100.413	<0.001	0.600
Block recall	55.800	<0.001	0.454
NKT	50.381	<0.001	0.429
WCST	14.255	<0.001	0.175
Stroop	12.114	<0.001	0.153

8.4.5 Age difference for biological constructs

An ANOVA was computed to examine the age differences in children's performance between each cohort for each biological task. The ANOVA revealed significant differences in performance between cohorts for all biological constructs (Table 8.10). Bonferroni *post*-hoc tests revealed that these differences were significant between every cohort for every biological construct and were all at $p<0.001$, except for performance on ecology between Year 2 and Year 5, which was at $p=0.015$.

Table 8.10. ANOVA for cohort differences in each biological construct

	F	*p*	partial η^2
Biodiversity	124.529	<0.001	0.650
Ecology	41.101	<0.001	0.380
Inheritance	25.372	<0.001	0.275
Evolution	49.631	<0.001	0.426

8.4.5.1 Differences in performance between biological construct

A repeated measures ANOVA was computed to assess whether children performed significantly better on any one area of biology than another at Time 1. This found there were significant differences in children's knowledge for biological constructs: $F(2.607)=222.810$, $p<0.001$, partial $\eta^2=0.621$. To explore these differences further, t-tests (Table 8.11) were conducted. These revealed that children performed significantly better on biodiversity than any other biological concept. There was no significant difference in children's performance on ecology and evolution knowledge, suggesting children perform very similarly on these two constructs. Inheritance was the weakest area of understanding; with children performing significantly better in all other areas in comparison.

Table 8.11. Descriptive statistics and paired samples t-tests of all biological constructs (df=136)

	Paired samples t-tests			
	Mean difference	Std. Deviation	Std. Error Mean difference	t
Biodiversity vs Ecology	1.108	0.728	0.0622	17.805**
Biodiversity vs Inheritance	1.614	0.929	0.079	20.345**
Biodiversity vs Evolution	1.109	0.724	0.061	17.932**
Ecology vs Inheritance	0.506	0.671	0.057	8.820**
Ecology vs Evolution	0.001	0.669	0.057	0.026
Inheritance vs Evolution	-0.505	0.781	0.067	-7.556**

*with Bonferroni correction the significance level is $p=0.008$ (two-tailed) **$p<0.001$

8.5 Partial correlations

Correlations were computed to see the extent to which scores for each of the biological constructs were associated with each other. These were done as a prelude to examining the relationship between the biological constructs and the general cognitive measures. Age was partialled out from these correlations, given the likely developmental trend. Table 8.12 shows all biological constructs are highly correlated with each other, though the values for inheritance were somewhat weaker:

Table 8.12. Partial correlations between all biological constructs, controlling for age. All df = 134

	Biodiversity	Ecology	Inheritance	Evolution
Biodiversity	1	0.600**	0.327**	0.558**
Ecology	0.600**	1	0.512**	0.550**
Inheritance	0.327**	0.512**	1	0.394**
Evolution	0.558**	0.550**	0.394**	1

**p*<0.05; ***p*<0.001 (two tailed)

Similar partial correlations (Table 8.13) were conducted with the measures for general cognitive ability to observe the extent to which these were related to each other. These correlations show that receptive language (BPVS) is significantly associated with all other general cognitive measures, apart from the WCST where there was no significant correlation. In fact the only significant association WCST had with another variable was with digit recall.

Table 8.13. Partial correlations for all general cognitive ability measures, controlling for age (df = 134).

	BPVS	Digit recall	b/w digit recall	Block recall	NKT	WCST	Stroop
BPVS	1	0.342**	0.252*	0.060*	0.236*	-0.144	-0.342**
Digit recall		1	0.332**	0.085	0.159	-0.238*	-0.11
B/W digit recall			1	0.137	0.290**	-0.115	-0.184*
Block recall				1	0.216*	-0.042	-0.293**
NKT					1	-0.071	-0.158
WCST						1	0.143
Stroop							1

**p*<0.05; ***p*<0.001 (two tailed)

Partial correlations between general cognitive measures and biological concept controlling for age at Time 1 were also conducted. These reveal that BPVS is the variable that is most associated with biological concepts. Evolution and biodiversity are the best predicted areas of understanding.

BPVS is significantly correlated with all biological concepts when controlling for age. Digit recall and backward digit recall are both significantly correlated with biodiversity and evolution. Block recall and WCST are not significantly correlated to any biological concept. The stroop task is significantly *negatively* correlated with ecology only, suggesting good inhibition skills are associated with greater performance on ecology. These results are presented in Table 8.14:

Table 8.14. Partial correlations between biological constructs and measures of general cognitive ability, controlling for age

Partial Correlations Time 1					
Control Variables		Biodiversity score	Ecology score	Inheritance score	Evolution score
Age at time of testing	BPVS	0.329**	0.376**	0.311**	0.338*
	Digit Recall	0.223*	0.083	0.058	0.265**
	Backward digit recall	0.290**	0.105	-0.017	0.212*
	Block recall	0.095	0.116	-0.090	0.098
	Number Knowledge score	0.166*	0.108	0.055	0.022
	WCST: perseverative errors	-0.119	-0.005	-0.105	-0.009
	Stroop: number of incongruent errors	-0.107	-0.151*	-0.071	-0.147

*p<0.05; **p<0.001

8.6. Parent demographic data

Some exploratory analyses were carried out to see the extent to which parent demographic data was related to children's scores at Time 1 for both biology and for general cognitive ability (Table 8.15).

Table 8.15. Cohort specific descriptive figures for parent demographics

	Reception (N=18)		Year 2 (N=36)		Year 5 (N=29)	
	Mean	SD	Mean	SD	Mean	SD
Number of adults	2.440	0.860	2.250	0.500	2.170	0.384
Occupation level of mother	1.060	1.890	2.060	2.229	2.550	2.501
Occupation level of father	3.170	1.920	2.860	1.726	3.620	1.935
Education level of father of child	1.50	1.340	1.500	1.320	1.970	1.210
Education level of mother of child	1.440	1.100	1.310	0.980	1.550	1.183
Number of older children at home	0.830	1.100	1.000	1.171	0.720	0.649
Number of younger children at home	0.670	0.910	0.640	0.639	0.720	1.032
Attendance of preschool	1.500	0.860	1.530	0.971	1.760	0.872
Free school meals (SES)	0.110	0.320	0.190	0.401	0.100	0.310
Native English	0.610	0.500	0.420	0.500	0.340	0.484

A multivariate ANOVA was conducted to see if there were any cohort differences in the demographic variables, excluding mothers' occupation level and number of children in the home given that these were two variables that were unlikely to vary with age. Findings revealed no significant differences between cohorts in any of the demographic variables, as can be seen in Table 8.16.

Table 8.16. Cohort differences in demographic variables

df=2	F	p	partial η^2
No. of adults	1.321	0.273	0.032
Free school meals (SES)	0.628	0.536	0.015
Attendance of preschool	1.434	0.244	0.035
Native English	1.647	0.199	0.040
Mothers' education	0.423	0.657	0.010
Fathers' education	1.235	0.296	0.030
Fathers' occupation	1.367	0.261	0.033

Table 8.17. Bivariate correlations between parent demographic variables

Control Variables (df = 80)	Number of adults in the home	Number of children in the home	Children who attended preschool yes/no	Free school meals	English native language	Education level of mother of child	Education level of father of child	Occupation level of mother	Occupation level of father
Number of adults in the home	1.000	-0.083	-0.071	-0.080	0.169	0.021	0.171	-0.056	0.030
Number of children in the home		1.000	0.009	0.250*	0.065	-0.272*	-0.204	-0.308*	-0.243*
Children who attended preschool yes/no			1.000	0-.100	-0.137	-0.022	0.073	0.093	0.083
Free school meals				1.000	0.118	-0.129	-0.316*	-0.069	-0.320*
English					1.000	-0.229*	-0.261*	-0.265*	-0.207

native language				
Education level of mother of child	1.000	0.573**	0.651**	0.530**
Education level of father of child		1.000	0.392**	0.769**
Occupation level of mother			1.000	0.459**
Occupation level of father				1.000

*p<0.05; **p<0.001

The parent demographic variables were also used in a correlational analysis, to see the extent to which parent variables were related to each other (Table 8.17). Overall mothers' and fathers' education level are significantly related to each other, as are their levels of occupation. Occupation and education levels of mothers and fathers are also associated with each other. Speaking English as a native language is associated with mothers' and fathers' education levels, and also mothers' occupation level. SES was only significantly associated with fathers' education and occupation levels, suggesting that these two variables might be a proxy for SES. Finally the number of children in the home is significantly associated with SES, fathers' education and occupation, and mothers' education.

These variables were then used in correlational analyses with the biological constructs, followed by a correlational analysis with the general cognitive ability measures, whilst controlling for age. The results of these analyses are shown below:

Table 8.18. Partial correlations between parent demographic data and each biological construct, controlling for age

Variables (df = 79)	Biodiversity	Ecology	Inheritance	Evolution
Number of adults in the home	0.029	0.048	-0.049	-0.036
Number of older children in the home	0.045	-0.087	-0.063	-0.077
Number of younger	0.136	0.057	0.074	0.145

children in the home				
Children who attended preschool yes/no	-0.112	-0.058	-0.146	-0.154
Free school meals	-0.055	-0.153	-0.013	-0.129
English native language	-0.246*	-0.227*	-0.214*	-0.368**
Education level of mother of child	0.028	0.188	0.093	0.287**
Education level of father of child	0.030	0.071	0.057	0.190*
Occupation level of mother	0.076	0.110	0.164	0.195
Occupation level of father	-0.009	0.049	0.111	0.078

*p<0.05; **p<0.001

The table shows that speaking English as a native language was significantly negatively correlated with each biological construct, but note, this is a function of how English language was scored and results actually imply that performance on the biological tasks was better if the children spoke English as native language. This is unsurprising given how language-based the biological task was, and how it would have required a child to have a good grasp of language to articulate themselves during the interview. Being a native speaker of English is also significantly associated with parent levels of education and occupation, suggesting those children who are native English speakers are more likely to have educated parents with higher levels of occupation.

Conversely, not being a native English speaker is significantly associated with free school meals (SES) implying that those who are from lower socio-economic backgrounds are less likely to be native English speakers.

The correlations also suggest that children's evolution knowledge is significantly associated with the education level of the mother and father, perhaps due to more exposure of evolutionary ideas.

Table 8.19. Partial correlations between parent demographic data and measures of general cognitive ability, controlling for age

Variables (df = 80)	BPVS	Digit recall	b/w Digit recall	Block recall	NKT	WCST	Stroop
Number of adults in the home	0.036	-0.093	-0.117	0.134	-0.126	-0.105	0.102
Number of older children in the home	-0.238*	0.023	-0.080	-0.203	0.016	-0.077	0.303*
Number of younger children in the home	-0.221*	-0.087	-0.059	-0.137	-0.029	-0.033	0.127
Children who attended preschool	-0.142	-0.023	0.045	0.124	-0.001	0.108	0.050

yes/no							
Free school meals	-0.160	-0.047	-0.121	-0.034	-0.041	-0.020	0.059
English native language	-0.481**	-0.193*	-0.341*	0.178*	-0.155	-0.159	-0.025
Education level of mother of child	0.272*	0.291*	0.035	0.138	0.105	-0.161	-0.201*
Education level of father of child	0.319*	0.114	0.147	0.016	0.158	-0.079	-0.203
Occupation level of mother	0.298*	0.322*	0.066	0.079	0.018	-0.129	-0.177
Occupation level of father	0.395**	0.187	0.194	-0.031	0.145	-0.080	-0.222

*p<0.05; **p<0.001

These correlations (Table 8.19) suggest that receptive language (BPVS) is significantly negatively associated with the number of children at home, implying that fewer children in the home leads to better receptive language skills, perhaps because of the opportunity for more varied speech with adults. Receptive language is also positively and significantly associated with the levels of education and occupation of both parents, suggesting high levels of parental education generally lead to better receptive vocabulary in the children, again this may be because of the opportunity for more sophisticated and varied speech.

Note that BPVS is also significantly negatively associated with having English as a native language, but given the way these data were scored this would imply that receptive language is better for children whose native language is English. This is unsurprising given the heavy reliance of the task on English language.

Verbal short-term memory (digit recall) is positively significantly associated with the occupation and education level of the mother. This would follow the same story with the associations with receptive language described above whereby better increased education may lead to better receptive language, and in turn, verbal short-term memory. Supporting this claim is the fact that verbal short-term memory and working memory are significantly associated with speaking English as a native language.

Finally, inhibition (Stroop task) is significantly negatively associated with the education level of the mother suggesting that better inhibition skills in the child is associated with how educated the child's mother is. Exactly why this may be the case is unclear but will be explored further on in this thesis.

CHAPTER 9: RESULTS TIME 2

9.1 Overview

This chapter outlines the analyses that were computed[4] after Phase Two (henceforth also known as Time 2) of the longitudinal study. The main focus of this chapter was to compare the differences between in children's performance for general cognitive ability test measures, and biological knowledge between Time 1 and Time 2, and to comment on any changes that occurred in this knowledge over time. Data coding took place during data collection at Phase Two.

Analyses were initially cross-sectional, concentrating on the group differences on all measures at Time 2. Following this, the longitudinal Time 1 to Time 2 changes in children's performance across all measures and between cohorts were explored. The effect of context, order of presentation, and gender are also presented. Parent demographic data and their relationship with general cognitive and biological measures at Time 2 are explored before considering the outcomes from data reduction techniques.

The main aim of this chapter was to establish the developmental change in general cognitive ability and in biological knowledge, and the trends that emerged. The subsequent chapter then goes on to explain the changes described here in this chapter.

[4] All data were analysed using IBM Statistics version 22, and Microsoft Excel version 2013

9.2 Descriptive analyses

Tests of normality were conducted on the data for biological knowledge and for general cognitive abilities, as in Time 1. The Kolmogorov-Smirnov and Shapiro-Wilk statistics are shown in Table 9.1:

Table 9.1. Tests of normality for measures of general cognitive ability and biological constructs at Time 2

Tests of normality	Cohort	Kolmogorov-Smirnov		Shapiro-Wilk	
		Statistic	df	Statistic	df
BPVS	Year 1	0.146*	43	0.900**	43
	Year 3	0.112	43	0.955	43
	Year 6	0.139*	43	0.969	43
	Whole sample	0.112**	129	0.958**	129
Digit Recall	Year 1	0.145*	43	0.959	43
	Year 3	0.187**	43	0.905*	43
	Year 6	0.113	43	0.969	43
	Whole sample	0.150**	129	0.966*	129
B/W Digit Recall	Year 1	0.217**	43	0.913*	43
	Year 3	0.166*	43	0.927*	43
	Year 6	0.176*	43	0.945*	43
	Whole sample	0.127**	129	0.964*	129
Block Recall	Year 1	0.152*	43	0.969	43
	Year 3	0.153*	43	0.922*	43
	Year 6	0.118	43	0.972	43
	Whole sample	0.136**	129	0.976*	129

Number knowledge	Year 1	0.131	43	0.961	43
	Year 3	0.132	43	0.963	43
	Year 6	0.169*	43	0.939*	43
	Whole sample	0.101*	129	0.968*	129
WCST	Year 1	0.171*	43	0.919*	43
	Year 3	0.208**	43	0.935*	43
	Year 6	0.288**	43	0.784**	43
	Whole sample	0.173**	129	0.916**	129
Stroop	Year 1	0.328**	43	0.662**	43
	Year 3	0.310**	43	0.604**	43
	Year 6	0.430**	43	0.532**	43
	Whole sample	0.329**	129	0.572**	129
Expressive Lang	Year 1	0.158*	43	0.955	43
	Year 3	0.095	43	0.984	43
	Year 6	0.133	43	0.948	43
	Whole sample	0.131**	129	0.953**	129
Biodiversity	Year 1	0.087	43	0.986	43
	Year 3	0.125	43	0.948*	43
	Year 6	0.083	43	0.976	43
	Whole sample	0.098*	129	0.975*	129
Ecology	Year 1	0.119	43	0.969	43
	Year 3	0.113	43	0.960	43
	Year 6	0.130	43	0.931*	43
	Whole sample	0.088*	129	0.958**	129
Inheritance	Year 1	0.259**	43	0.747**	43
	Year 3	0.212**	43	0.790**	43

	Year 6	0.140*	43	0.872**	43
	Whole sample	0.198**	129	0.783**	129
Evolution	Year 1	0.133	43	0.947*	43
	Year 3	0.226**	43	0.794**	43
	Year 6	0.137*	43	0.924*	43
	Whole sample	0.178**	129	0.783**	129

*p<0.05; **p<0.001

The tests of normality suggest that the majority of data are not normally distributed, with some variation between cohorts as there was at Time 1, aside from expressive language where the data for all cohorts are normally distributed (but not across the whole sample). These significant values tend to be concentrated among the younger cohorts, just like Time 1, suggesting that they are largely a consequence of floor effects. Likewise, the statistics for kurtosis and skewness across and between cohorts (Table 9.2) suggest very few extreme departures from the mean, and in any case, the sample size for these analyses is sufficient to proceed with parametric analyses.

The variables for biological constructs (Table 9.2) and general cognitive ability measures (Table 9.3) were analysed to observe children's performance between cohorts.

Table 9.2. Mean and standard deviation values for biological constructs, across cohorts at Time 1 and Time 2

		Time 1				Time 2				
		Mean	SD	Skewness	Kurtosis		Mean	SD	Skewness	Kurtosis
Biodiversity	R'tion	1.641	0.507	0.198	-0.396	Year 1	2.578	0.813	0.156	0.663
	Year 2	2.763	0.725	0.337	-0.177	Year 3	3.687	1.054	0.767	0.307
	Year 5	3.804	0.745	-0.018	-0.632	Year 6	5.163	1.012	-0.052	-0.695
	Whole sample	2.719	1.110	0.267	-0.802		3.809	1.431	0.321	-0.641
Ecology	R'tion	1.038	0.404	0.963	1.212	Year 1	1.349	0.473	0.097	-0.340
	Year 2	1.527	0.615	0.908	0.821	Year 3	2.014	0.578	0.479	-0.295
	Year 5	2.278	0.898	0.815	0.406	Year 6	2.845	0.816	0.668	-0.497
	Whole sample	1.607	0.839	1.224	1.616		2.069	0.883	0.751	0.465
Inheritance	R'tion	0.609	0.432	0.506	-1.111	Year	0.539	0.478	1.648	1.645

						1				
	Year 2	1.161	0.618	0.841	0.389	Year 3	0.936	0.811	1.897	4.328
	Year 5	1.554	0.842	0.853	0.184	Year 6	1.710	1.263	1.299	1.399
	Whole sample	1.1004	0.755	1.097	1.258		1.061	1.03	1.932	4.256
Evolution	R'tion	0.896	0.500	0.583	0.090	Year 1	1.567	0.678	0.550	1.768
	Year 2	1.643	0.796	1.258	2.831	Year 3	2.328	0.965	2.109	6.034
	Year 5	2.324	0.773	0.705	0.746	Year 6	3.634	1.245	0.402	-1.160
	Whole sample	1.610	0.911	0.761	0.605		2.510	1.304	1.012	0.551

Table 9.3. Mean and standard deviation values for measures of general cognitive ability across cohorts

		Time 1				Year 2				
		Mean	SD	skewness	kurtosis		Mean	SD	Skewness	Kurtosis
BPVS	R'tion	63.230	17.533	-0.441	0.307	Year 1	79.420	12.726	-1.367	3.090
	Year 2	86.560	13.910	-0.710	1.299	Year 3	100.560	18.183	0.421	-0.568
	Year 5	121.360	16.472	-0.002	-0.741	Year 6	131.050	13.459	-0.473	-0.132
	Whole sample	89.430	29.459				103.670	25.955	0.095	-0.853
Digit recall trials	R'tion	22.830	3.867	0.052	-0.424	Year 1	24.880	3.500	-0.171	-0.173
	Year 2	27.870	4.556	0.364	0.549	Year 3	27.000	4.117	0.680	1.804
	Year 5	32.620	5.462	0.859	0.338	Year 6	29.420	4.289	0.209	-0.664
	Whole sample	27.670	6.135	0.556	0.068		27.10	4.368	0.388	0.331
B/W digit recall trials	R'tion	5.470	1.544	0.007	1.634	Year 1	6.260	3.787	-0.131	-0.525
	Year 2	9.980	3.625	0.915	0.445	Year 3	11.470	3.838	0.569	-0.603
	Year 5	14.780	3.855	0.204	-1.050	Year 6	14.950	5.318	0.394	-0.861
	Whole sample	10.010	4.960	0.652	-0.529		10.38	5.629	0.407	0.064
Block recall	R'tion	16.890	3.908	-0.487	1.672	Year 1	21.600	3.424	0.242	-0.044

trials	Year 2	20.780	3.437	-0.324	-0.297	Year 3	23.510	4.183	0.819	-0.028
	Year 5	25.220	3.971	-0.140	0.707	Year 6	26.510	3.615	0.141	0.769
	Whole sample	20.890	5.097	-0.068	0.223		23.880	4.241	0.366	-0.184
Number knowledge	R'tion	77.997	6.775	-0.543	-0.387	Year 1	77.822	5.370	0.110	1.184
	Year 2	78.211	8.922	-0.772	1.561	Year 3	84.231	7.045	0.273	-0.183
	Year 5	91.599	6.230	-0.924	0.112	Year 6	92.967	4.585	-0.221	-0.964
	Whole sample	82.570	9.740	-0.354	0.077		85.007	8.455	-0.007	-0.844
WCST	R'tion	14.810	7.503	1.176	0.606	Year 1	8.740	6.762	0.816	-0.026
	Year 2	10.440	4.256	1.730	5.797	Year 3	9.510	3.418	-0.266	1.495
	Year 5	9.220	2.875	0.331	1.810	Year 6	9.290	2.030	-1.791	3.544
	Whole sample	11.580	5.784	1.891	4.217		9.180	4.505	0.616	2.004
Stroop task	R'tion	2.430	2.717	1.960	4.666	Year 1	1.530	2.585	2.208	5.662
	Year 2	1.390	1.919	1.443	1.268	Year 3	1.090	1.962	2.543	6.432
	Year 5	0.420	0.543	0.768	-0.536	Year 6	0.350	0.720	2.570	7.091
	Whole sample	1.430	2.114	2.480	8.037		0.990	1.967	2.916	9.992
Expressive language						Year 1	17.120	4.288	-0.486	-0.589
						Year 3	27.72	7.129	0.162	-0.430

		Year 6	42.51	8.028	-0.293	-1.008
		Whole sample	29.12	12.376	0.401	-0.891

9.3 Significance testing at Time 2

9.3.1 Gender

There were no significant differences in gender for either the general cognitive ability measures, or the biological constructs; hence gender differences are not examined further at Time 2.

9.3.2 Order of presentation of context

There were no significant differences in children's performance for the order in which the two contexts were presented across the overall sample, except for evolution: $F(127)=4.176$, $p=0.022$, one-tailed, where children generally performed better in the second context (M=2.740, SD=1.443) than the first (M=2.277, SD=1.108). This is in contrast to Time 1, where order effects were only seen for inheritance contexts. When this was examined further, the only cohort displaying significant differences in the order of presentation of contexts was Cohort 2 (Year 3) where children performed significantly better on the second context presented (M=2.585, SD=1.231) than the first (M=2.104, SD=0.597): $F(41)=2.766$, $p=0.05$, one-tailed.

9.3.3 Context

A mixed ANOVA was conducted to investigate if there were any contextual differences in children's performance by cohort. The analysis revealed significant differences and these were explored further by *post-hoc* t-tests (Bonferroni corrected significant p value =0.006). There were significant differences in contextual knowledge across children only for biodiversity and ecology contexts. For biodiversity, children held significantly more knowledge in the context of a savannah (M=1.548, SD=0.659) than the pond context (M=1.482, SD=0.616): $t(128)=-2.327$, $p=0.011$, one-tailed. For ecology, children held significantly more knowledge in the context of pond (M=1.107, SD=0.465), than savannah (M=0.963, SD=0.449): $t(128)=6.840$, $p<0.001$, one-tailed. There were no contextual differences for inheritance and evolution concepts.

Contextual differences between cohorts were also analysed (Table 9.4; Bonferroni corrected significant p value =0.004). For children in Year 1, there was only a significant contextual difference in children's performance for ecological concepts where children performed better in the pond context (M=0.737, SD=0.256) than the savannah context (M=0.612, SD=0.257): $t(42)=4.140$, $p<0.001$, one-tailed).

For children in Year 3, there were significant contextual differences in performance for ecology concepts. Children performed better in the pond context (M=1.080, SD=0.334) than the savannah context (M=0.935, SD=0.299): $t(42)=3.619$, $p<0.001$, one-tailed.

For children in Year 6, there were only contextual differences between ecological concepts where children performed better in the pond context (M=1.504, SD=0.423), than in the savannah context (M=1.342, SD=0.432): $t(42)=-0.206$, $p<0.001$, one-tailed.

Table 9.4. Contextual differences in children's performance across biological constructs by cohort

		Pond		Savannah		t (df=42)	Sig (one-tailed)
		Mean	SD	Mean	SD		
Biodiversity	Cohort 1	0.973	0.386	1.023	0.385	1.223	0.114
	Cohort 2	1.436	0.485	1.480	0.528	1.062	0.147
	Cohort 3	2.040	0.429	2.142	0.497	1.660	0.052
Ecology	Cohort 1	0.738	0.256	0.612	0.257	4.140	0.001
	Cohort 2	1.080	0.334	0.935	0.299	3.619	0.001
	Cohort 3	1.504	0.423	1.342	0.432	4.149	0.001
Inheritance	Cohort 1	0.930	0.225	0.105	0.279	-0.253	0.400
	Cohort 2	0.244	0.384	0.198	0.396	0.942	0.176
	Cohort 3	0.477	0.556	0.488	0.668	-0.206	0.031
Evolution	Cohort 1	0.686	0.296	0.707	0.287	-0.606	0.274
	Cohort 2	0.974	0.331	0.912	0.353	2.414	0.010
	Cohort 3	1.386	0.365	1.354	0.436	1.022	0.157

** Bonferroni corrected significance level *p*=0.004 (one-tailed)

These contextual differences that demonstrate, in general, a preference for the pond context were not present at Time 1. At Time 1 only a marginally significant effect of context was seen for inheritance where children in Cohort 2 preferred the savannah context. Reasons for why the prevalent preference for the pond context is present at Time 2 could vary. It may be that children encounter pond contexts more than savannah contexts in everyday life, and earlier exposure to these contexts at Time 1 made children more sensitised to acquire new knowledge within the same contexts they are repeatedly exposed to. Also, teachers and school trips regularly focus on pond environments as opposed to savannahs, and this is also likely to contribute to the differences in contexts observed at Time 2.

9.3.4 Age differences in general cognitive abilities

A factorial ANOVA was conducted to investigate whether there were any significant age differences in children's performance for any of the measures for general cognitive ability. The ANOVA revealed a main effect of year group (see Table 9.6) for all measures except the number of perseverative errors in the WCST, where children performed similarly across cohorts (see Table 9.5). Bonferroni *post-hoc* one-tailed t-tests were computed to explore the age differences in further detail. The majority of these tests revealed that there were significant differences at the $p<0.001$ level between all age groups on BPVS, digit recall, backwards digit recall, block recall, NKT, and expressive language where children in Year 6 performed significantly better than children in Year 3, who performed significantly better than children in Year 1. However, there were no significant differences between age groups in the WCST (all at $p=1$), where children in Year 1 (M=8.744, SD=6.762) made the fewest perseverative errors, followed by Year 6 (M=9.287, SD=2.030), then Year 3 (M=9.512, SD=3.418). This pattern would imply children in Year 1 perform the best on this task, however, the pattern could also reflect the fact that children in the youngest cohort simply did not understand the task and while they made fewer perseverative errors, they also made the fewest number of rule change detections, suggesting poor performance overall, as shown in Table 9.5 below.

Table 9.5. Number of successful rule changes made in the WCST at Time 2

		Year 1	Year 3	Year 6
Rule changes detected	Mean	6.360	9.260	11.070
	SD	3.370	2.735	1.724

For the Stroop task, the only significant difference was between Year 1 (M=1.535, SD=2.585) and Year 6 (M=0.349, SD=0.720) at *p*=0.015 where children in Year 6 made significantly fewer incongruent errors than children in Year 1. Whereas children in Year 3 (M=1.093, SD=1.962) did not perform significantly differently to either children in Year 1 (*p*=0.432) or Year 6 (*p*=0.112).

Table 9.6. ANOVA results for age differences in general cognitive measures

df=2	F	*p*	partial η^2
BPVS	128.982**	<0.001	0.672
Digit recall	13.956**	<0.001	0.181
B/W digit recall	43.094**	<0.001	0.406
Block recall	18.67**	<0.001	0.229

NKT	74.898**	<0.001	0.543
WCST	0.326	0.361	0.005
Stroop task	4.195*	0.009	0.062
Expressive language	157.016**	<0.001	0.714

**p*<0.05; ***p*<0.001 (one-tailed)

9.3.5 Age differences in biological constructs

A factorial ANOVA was computed to examine the age differences in children's knowledge for each biological construct (biodiversity, ecology, inheritance, and evolution). This analysis revealed a main effect of year group for each construct (Table 9.7). Bonferroni *post-hoc* tests further illustrated that children in Year 6 performed significantly better than children in Year 3, who in turn performed significantly better than children in Year 1 across all biological constructs, all at $p<0.001$ (see Table 9.2 for Mean values). These findings generally echo those established in Time 1, where the same results were found.

Table 9.7. ANOVA results for age differences in biological constructs

(df=2)	F	partial η^2
Biodiversity	77.603**	0.552
Ecology	59.394**	0.485
Inheritance	18.398**	0.226
Evolution	47.956**	0.432

***p*<0.001 (one-tailed)

9.3.6 Differences in performance between biological constructs

In order to investigate whether there were any differences in children's performance for one biological construct over another at Time 2, an ANOVA was computed. This showed that there was a main effect of biological concept on children's performance: $F(3,375)=619.858$, $p<0.001$, partial $\eta^2=0.832$. *Post-hoc* t-tests were computed to examine this effect further (Table 9.8). These analyses revealed children performed significantly better on biodiversity (M=3.809, SD=1.431) than ecology (M=2.069, SD=0.883), inheritance (M=1.061, SD=1.026), or evolution (M=2.510, SD=1.303), all at $p<0.001$. Children also performed significantly better on ecology than inheritance, but significantly worse in comparison to evolution (both at $p<0.001$). Lastly, children performed significantly better on evolution than inheritance, $p<0.001$). These results are on the whole similar to those at Time 1 where children's strongest area of performance was biodiversity, and weakest area was inheritance. However at Time 1, there was no significant difference between ecology and evolution, whereas by Time 2, children are performing significantly better in evolution concepts, than ecology.

Table 9.8. Post-hoc analyses to examine differences in performance for biological constructs

Paired samples t-tests (one-tailed) df=128				
	Mean difference	Std. Deviation	Std. Error Mean difference	t
Biodiversity vs Ecology	1.739	0.878	0.077	22.492**
Biodiversity vs Inheritance	2.747	1.197	0.105	26.064**
Biodiversity vs Evolution	1.299	0.859	0.075	17.151**

Ecology vs Inheritance	1.008	0.937	0.082	12.221**
Ecology vs Evolution	-0.440	0.075	0.075	-5.843**
Inheritance vs Evolution	-1.448	0.089	0.089	-16.344**

**Bonferroni corrected significance level = *p*=0.008)

9.4 Significance testing-Time 1 to Time 2 changes

9.4.1 General cognitive abilities

A series of two-way mixed ANOVAs were conducted for each of the general cognitive abilities to see if there was any effect of time point or year group on children's scores.

9.4.1.1 BPVS

Regarding the BPVS scores, there was a significant main effect of time point $F(1,126)=173.427$, $p<0.001$, partial $\eta^2=0.579$, and a significant main effect of year group, $F(2,126)=153.302$, $p<.001$, partial $\eta^2=0.709$. This suggests that that children across all cohorts are performing significantly better at Time 2 than Time 1, and that at both time points children in Cohort 3 perform significantly better than children in Cohort 2, who in turn perform significantly better than children in Cohort 1, as shown by *post-hoc* t-tests in Table 9.9. The significant interaction between time point and year group, $F(2,126)=4.492$, $p=0.013$, partial $\eta^2=0.067$ also indicates that although children's BPVS scores are significantly different at Time 2 from Time 1, these differences in scores are different depending on year group. Generally, the biggest change from Time 1 to Time 2 is with Cohort 1, followed by Cohort 2, and the smallest change is with Cohort 3.

9.4.1.2 Digit recall

For digit recall, there was no significant main effect of time point ($F(1,126)=3.272$, $p=0.073$, partial $\eta^2=0.025$) suggesting that children's scores for digit recall at Time 2 are not significantly different from their scores at Time 1. There was a significant main effect of year group: $F(1,126)=34.726$, $p<0.001$, partial $\eta^2=0.355$. *Post-hoc* t-tests (Table 9.9) revealed there were significant Time 1 to Time 2 changes only for Cohort 1 and Cohort 3, but not for Cohort 2 who performed similarly across the two time points.

The two-way ANOVA also revealed an interaction between time point and year group: $F(1,126)=16.145$, $p<0.001$, partial $\eta^2=0.204$. It seems as though Cohort 3, the oldest cohort, significantly decline in their digit recall performance. This may imply that children have reached span by this age, particularly because Cohort 2 also decline in their performance by Time 2, although not significantly from Time 1. Conversely, Cohort 1, the youngest cohort, significantly improve in their performance over time points.

9.4.1.3 Backwards digit recall

Backward digit recall revealed a significant main effect of time point ($F(1,126)=115.004$, $p<0.001$, partial $\eta^2=0.477$) and year group ($F(1,126)=40.037$, $p<0.001$, partial $\eta^2=0.389$), suggesting children were performing significantly differently at the two time points as well as between cohorts. Further *post-hoc* t-tests (Table 9.9) revealed this significant difference was only for Cohort 2 whose performance for backwards digit recall significantly improved at Time 2. Children in Cohorts 1 and 3 also improved in their performance, although not significantly. A lack of significant improvement for Cohort 3 may be because children have reached span, as with the digit recall task. With regards to Cohort 1 however, it may imply that the task is still a little difficult for the majority of children, and their performance overall is significantly lower than children in Cohort 2 and 3 as shown in section 9.3.4. Finally, the two-way ANOVA revealed a significant interaction between year group and time point: $F(1,126)=14.437$, $p<0.001$, partial $\eta^2=0.186$.

9.4.1.4 Block recall

Results for block recall showed a significant main effect of time point ($F(1,126)=115.044$, $p<0.001$, partial $\eta^2=0.477$) and year group ($F(1,126)=40.037$, $p<0.001$, partial $\eta^2=0.389$). *Post-hoc* t–tests (Table 9.9) indicated that children across all cohorts performed significantly better at Time 2 than Time 1, and that children in Cohort 3 performed significantly better than children in Cohort 2, who in turn performed significantly better than children in Cohort 1. This is illustrated by the significant interaction between year group and time point too, $F(1,126)=14.437$, $p<0.001$, partial $\eta^2=0.186$, suggesting that although scores were significantly different between time points, the way in which they differed varied according to the year group of the children. Generally children in Cohort 1 had the biggest improvement in their performance, followed by Cohort 2, and then Cohort 1.

9.4.1.5 Number Knowledge

Results for number knowledge revealed that there was a significant main effect of time point (F(1,126)=13.133, p<0.001, partial η^2 =0.094), year group (F(1,126)=79.453, p<0.001, partial η^2 =0.558), and a significant interaction between time point and year group (F(1,126)=8.978, p<0.001, partial η^2 =0.125. *Post-hoc* t-tests (Table 9.9) revealed that only children in Cohort 2 were performing significantly better at Time 2 than Time 1, whereas children in Cohort 3 were performing better but not significantly. Conversely children in Cohort 1 were performing worse, although this change was not significant. It may be that the older children in Cohort 3 found the task relatively easy and many were performing at ceiling, contributing to lack of significant change over time. With regards to the children in youngest cohort, it may be that children failed to engage with the task as well as children in the other cohorts because their overall maths ability may still be quite low. This would mean that they would be unlikely to progress to harder questions in the task, and do poorly overall.

9.4.1.6 Perseverative errors–WCST

With regards to the number of perseverative errors children make in the WCST, the two-way repeated measures ANOVA revealed a significant main effect of time point (F(1,126)=13.190, p<0.001, partial η^2 =0.095), a significant main effect of year group (F(1,126)=6.572, p<0.001, partial η^2 =0.094), and a significant interaction between time point and year group (F(1,126)=8.906, p<0.001, partial η^2 =0.124). Further *post-hoc* analysis illustrated there was only a significant change in performance from Time 1 to Time 2 for the youngest cohort, who made fewer perseverative errors than children in both Cohorts 2 and 3 (Table 9.9). However as explained earlier in section 9.3.4, this may reflect poor performance overall as these children detect the fewest number of rule changes overall (Table 9.5). Time 1 and Time 2 scores for the oldest cohort (Year 5 & 6) are very similar and not significantly different so this may depict a ceiling effect, or more likely, a failure to engage with the task, given its simplicity. Cohort 2 seem to be performing as expected, by making fewer preservative errors at Time 2 than at Time 1, however this change is not significant.

9.4.1.7 Incongruent errors–Stroop task

Finally, the ANOVA for the number of incongruent errors made during the Stroop task also revealed a significant main effect of time point (F(1,126)=7.029, p=0.009, partial η^2 =0.053), and year group (F(1,126)=8.977, p<0.001, partial η^2 =0.126), but no significant interaction between time point and year group (F(2, 125) = 1.822, p=0.166, partial η^2 =0.028). *Post-hoc* analysis highlighted that all cohorts performed better at the task at Time 2, but that the change in scores from Time 1 was only significant for Cohort 1, hence the lack of a significant interaction between year group and time point. Generally children in Cohort 3 performed better (but not significantly better) than children in Cohort 2, who in turn performed better (but not significantly better) than children in Cohort 1. Cohort 3 only performed significantly better than children in Cohort 1 (Table 9.9). It seems that children in the middle cohort were performing similarly to those in cohorts above and below it.

9.4.1.8 Summary

Overall, for Cohort 1 (Year 1/Reception children), the Time 1 to Time 2 changes are all significantly different after Bonferroni corrections (p=0.007), except changes in backwards digit span, t(42)=-1.619, p=0.50, one-tailed, number knowledge t(42)=.436, p=.0.332, one-tailed, and Stroop task t(42)=2.121, p=0.020, one-tailed, where a small and non-significant improvement is made.

For Cohort 2 (Year 2/3) children, the Time 1 to Time 2 changes are all significantly different except for: changes in digit span where children perform similarly across time points, t(42)=1.436, p=0.079, one-tailed, changes in WCST where the same pattern is seen, t(42)=1.100, p=0.139, one-tailed, and again for the Stroop task, where children are performing better at Time 1 than Time 2, t(42)=1.415, p=0.082, one-tailed.

For the oldest cohort (Year 5/6) the Time 1 to Time 2 changes follow a similar pattern, with all performance being significantly better at Time 2 except for backwards digit span where performance at Time 2 was not significantly better, t(42)=-0.614, p=0.271, one-tailed, block recall, t(42)=-2.401, p=0.011, one-tailed, number knowledge, t(42)=-1.850, p=0.035, one-tailed, WCST, t(42)=-.111, p=0.456, one-tailed, and Stroop task, t(42)=.662, p=0.256, one-tailed, where there was also a non-significant improvement in scores at Time 2. There are generally fewer significant Time 1 to Time 2 changes in Year 5/6 cohort than any other cohort of children, presumably because this cohort is reaching ceiling or span on many of the tasks.

Table 9.9. Differences in general cognitive ability performance across Time 1 to Time 2 between cohorts

(df=42)		year 1		year 2			
		Mean	SD		Mean	SD	*t* value
BPVS	reception	64.240	17.690	year 1	79.880	12.770	-8.429**
	year 2	86.880	13.460	year 3	100.560	18.180	-7.941**
	year 5	121.980	16.500	year 6	131.050	13.460	-6.280**
digit recall	reception	22.800	3.560	year 1	25.070	3.410	-3.905**
	year 2	27.950	4.640	year 3	27.000	4.120	1.436
	year 5	32.330	5.390	year 6	29.420	4.290	4.482**
B/W digit recall	reception	5.390	1.430	year 1	6.390	3.790	1.619
	year 2	9.980	3.670	year 3	11.470	3.840	-3.027*
	year 5	14.560	3.790	year 6	14.950	5.320	-0.614
block recall	reception	17.000	4.090	year 1	21.730	3.430	-10.134**
	year 2	20.700	3.500	year 3	23.510	4.180	-6.246**
	year 5	25.330	3.980	year 6	26.510	3.610	-2.401
Number knowledge	reception	78.330	6.920	year 1	77.820	5.370	0.436
	year 2	77.940	9.000	year 3	84.230	7.050	-4.433**
	year 5	91.440	6.240	year 6	92.970	4.580	-1.850
WCST perseverative errors	reception	14.440	7.400	year 1	8.540	6.680	3.675**
	year 2	10.470	4.310	year 3	9.510	3.420	1.100
	year 5	9.230	2.930	year 6	9.290	2.030	-0.111
Stroop incongruent errors	reception	2.410	2.870	year 1	1.560	2.640	2.121
	year 2	1.440	1.920	year 3	1.090	1.960	1.415
	year 5	0.440	0.550	year 6	0.350	0.720	0.662

*Bonferroni corrected level of significance: *p*=0.007 **, one-tailed; N=43 for each cohort

9.4.2 Biological concepts

A four x two mixed ANOVA was conducted to examine to relationships between biological concepts and time point between cohorts (Table 9.10). The effects all violate the assumption of sphericity, therefore Greenhouse-Geisser statistics will be reported. There was a significant main effect of biological concept, $F(3,375)=619.858$, $p<0.001$, partial $\eta^2=0.832$, hence regardless of time of testing and year group, children's knowledge varied according to concept: children demonstrated more knowledge on biodiversity than evolution, $F(1,125)=624.340$, $p<.001$, $\eta^2=0.833$, more knowledge of evolution than ecology, $F(1,125)=22.186$, $p<0.001$, partial $\eta^2=0.151$, and finally more knowledge of evolution over inheritance, $F(1,125)=340.569$, $p<0.001$, partial $\eta^2=0.732$.

There is also a significant effect of time point, $F(1,125)=123.612$, $p<0.001$, partial $\eta^2=0.497$, hence irrespectively of biological concepts and year group, the time of testing significantly effects children's knowledge of the concept i.e. it suggests that the more time and experience children had around a particular topic, the more knowledge they held about that particular topic regardless of what that topic was.

Further *post-hoc* analyses revealed that in general, all results for biological concepts are depicting a developmental trend with children improving in biological knowledge across all four areas of biology, but these improvements were not consistent across inheritance concepts. Reception children improve significantly from Time 1 to Time 2 on all areas of biology except inheritance, *t*(42)=-0.660, *p*=0.257, one-tailed, where there is a slight decline in progress. The same is also true for children in the Year 2/3 cohort, *t*(42)=0.613, *p*=0.272, one-tailed, where children perform slightly worse at Time 2 than Time 1. Finally for the Year 5/6 cohort, an improvement is seen is all areas of biology, including inheritance, although for the latter the improvement is non-significant, *t*(42)=-0.685, *p*=0.249, one-tailed.

Table 9.10. Differences in performance of biological knowledge task across Time 1 and Time 2 for all cohorts

		Time 1			Time 2		
		Mean	SD		Mean	SD	*t* value
Biodiversity	R'tion	1.600	0.500	Year 1	2.580	0.810	-8.502**
	Year 2	2.800	0.700	Year 3	3.710	1.060	-6.427**
	Year 5	3.800	0.800	Year 6	5.160	1.010	-10.48**
Ecology	R'tion	1.010	0.400	Year 1	1.350	0.470	-4.711**
	Year 2	1.520	0.630	Year 3	2.020	0.580	-4.494**
	Year 5	2.330	0.890	Year 6	2.850	0.820	-3.611**
Inheritance	R'tion	0.600	0.430	Year 1	0.540	0.480	0.660
	Year 2	1.140	0.580	Year 3	0.950	0.820	1.404
	Year 5	1.560	0.860	Year 6	1.710	1.260	-0.685
Evolution	R'tion	0.850	0.490	Year 1	1.570	0.680	-6.109**
	Year 2	1.650	0.800	Year 3	2.340	0.970	-5.105**
	Year 5	2.330	0.790	Year 6	3.630	1.250	-6.361**

** Bonferroni corrected level of significance *p*=0.013, one-tailed; N=43 for each cohort

In terms of between-subjects effects, Levene's test indicates that variances are not homogenous. However, Levene's test provides a more conservative measure likely to lead towards a type two error, but the fact that a main effect still exists given this conservative test, and the large sample size of this study, it can be confidently assumed that the effects seen are reliable. Analysis revealed there was a main effect of year group, $F(2,125)=104.925$, $p<0.001$, partial $\eta^2=0.627$, suggesting that children's scores significantly differed across year groups regardless of the type of construct it was, or the time point at which they were tested. Generally children in Cohort 3 performed significantly better than children in Cohort 2, who in turn performed significantly better than children in Cohort 1, as described earlier in section 9.3.5.

The mixed ANOVA also showed a significant two-way interaction between construct and year, $F(6,375)=21.023$, $p<0.001$, partial $\eta^2=0.252$, indicating that although children's scores for a particular construct were affected by the type of construct it was (biodiversity, ecology etc.), the way in which scores were affected by concept differed between the year groups as shown by Figure 9.1 below. This plot illustrates the pattern of change in performance across time points for each cohort, and seems to suggest that the interaction reflects the growing size of gap between the constructs with age:

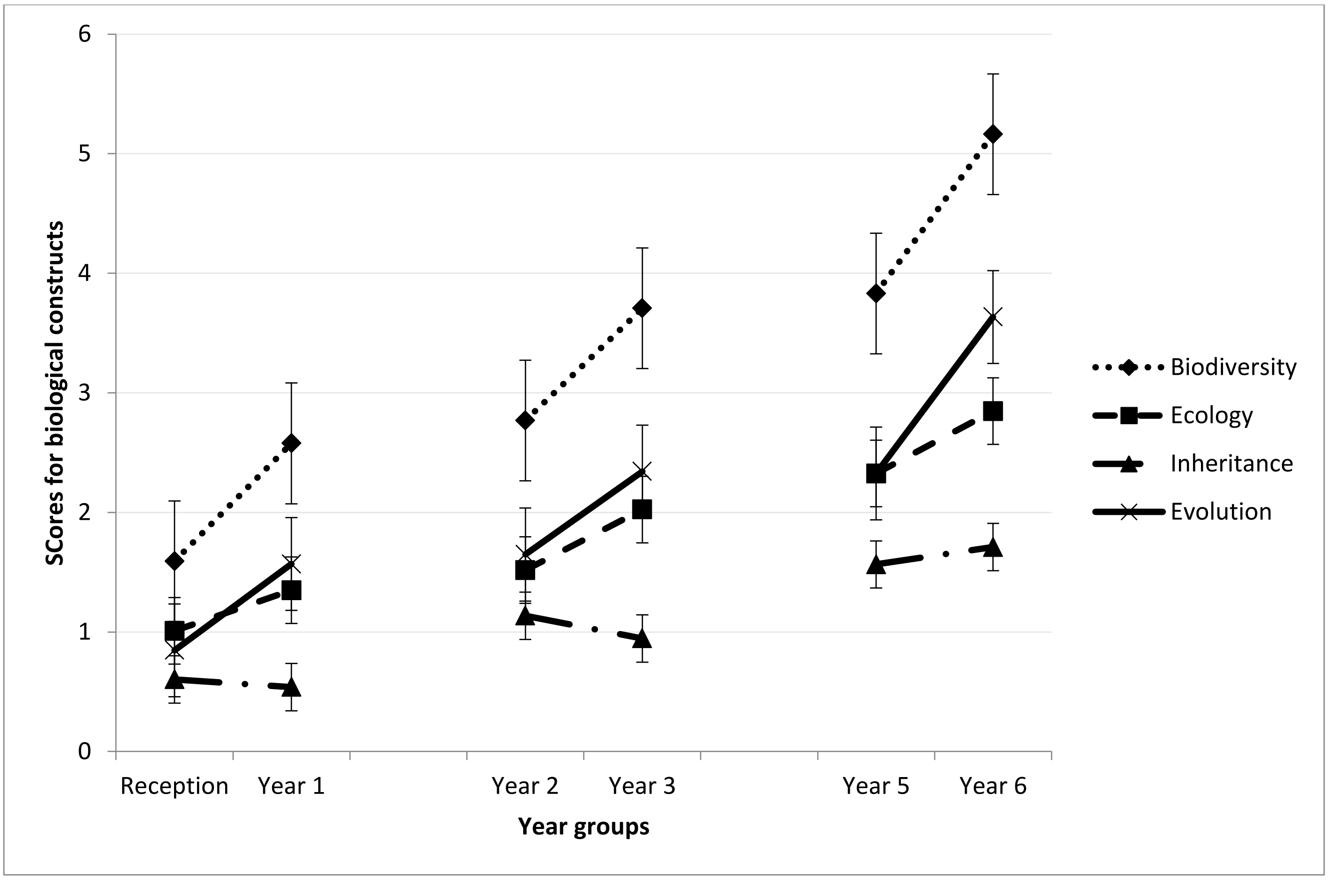


Figure 9.1. Graph illustrating the pattern of Time 1 to Time 2 changes across cohorts for performance on each biological construct

The construct and time point interaction was also significant, $F(3,375)=51.759$, $p<0.001$, partial η^2 =0.293, indicating that although children's scores differed by the type of construct, the way in which these scores differed by construct was different at the different time points, generally with biodiversity and evolution concepts making significant progress, followed by some progress with ecology concepts, and very little with inheritance concepts, where scores get further apart from Time 1 to Time 2.

Figure 9.1 suggests children generally hold the most knowledge about biodiversity initially, followed by evolution and ecology where children perform similarly, while knowledge on inheritance is generally low. But at Time 2, while biodiversity continues to have high scores, children increase in their evolutionary concepts at a similar rate to that of biodiversity, surpassing ecology, followed by inheritance, which remains the concept with the lowest scores throughout and actually declines slightly by Time 2.

There is very little carry over from Time 1 to Time 2 for inheritance, especially when looking at the earlier plots from the mixed ANOVA that help to visualise the relationships. The mean scores for inheritance are consistently low, with ecology and evolution at roughly the same level at Time 1 (but evolution generally overtakes ecology), while biodiversity means are consistently highest. The fact that inheritance concepts are lowest may suggest that it is being measured in a different way; the context-specific alphas for inheritance show very low reliability because there are very few context-specific elements, unlike the other constructs. While this makes sense, given that many of the questions around inheritance are on reproduction and are thus not context-specific, there is an argument that therefore inheritance is not being treated in quite the same way as the other three biological constructs.

An important note to consider is that Figure 9.1 makes apparent the implication that a lack of progress between cohorts for all biological constructs (except inheritance) is unlikely to be true. Hence, this would suggest that the measurement task itself to some extent is inducing progress over time point. This is a striking finding. For example, scores for biodiversity concepts at Time 2 for Cohort 1 are equivalent to scores for the same concepts at Time 1 for Cohort 2, and yet there is a difference of a whole school year in between. This suggests that the very act of undergoing testing procedures and taking part in the biological task appears to be boosting the performance of children at Time 2 within a cohort, to the same level of performance at Time 1 of children in a full academic year later. Furthermore, despite the fact that the trajectories for each biological construct are somewhat different, the effect of testing appears to be very uniform for the majority of biological concepts, bar inheritance, (which seems to have a different trajectory altogether) and therefore cannot be a simple artefact of the data. The importance of this finding will be discussed further in Chapter 11. Analysis also revealed a significant interaction between time point and year group (Figure 9.2) demonstrating at although there are significant changes from Time 1 to Time 2, the rate of change is influenced by the year group the child is in, $F(2,125)=4.265$, $p=0.016$, partial $\eta^2=0.064$, where Cohort 3 consistently demonstrate most knowledge and make more progress from Time point 1 to 2, whereas Cohort 2 and Cohort 1 make similar levels of progress across the two time points.

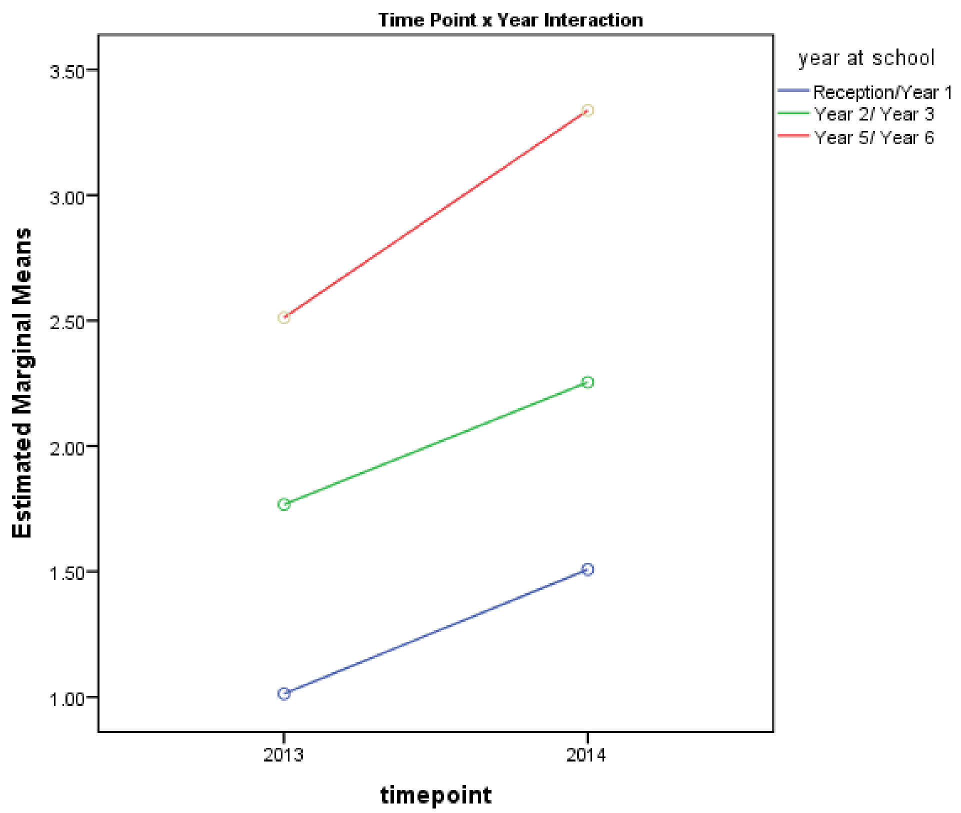


Figure 9.2. Graph illustrating two-way interaction between time point and cohort

There is no significant three-way interaction between concept, time point, and year, $F(6,375)=1.275$, $p=0.268$, partial $\eta^2=0.020$.

9.5 Parent demographic data

Some exploratory analyses were carried out to see the extent to which parent demographic data were related to children's scores at Time 2 for both biological constructs, and for general cognitive abilities. Using the questionnaire given to parents to complete, the same 10 variables selected at Time 1 were used for these analyses.

A partial correlational analysis (Table 9.11) was computed to see the extent to which parent variables were related to knowledge on the biological constructs at Time 2, whilst controlling for age. These illustrate that the only variables significantly related to the four biological constructs are: number of adults in the home, English spoken as native language, mothers' education level, and parental occupation levels. Aside from the number of adults in the home which is only significantly associated with children's biodiversity level at Time 2, the remaining variables are all significantly associated with biodiversity, ecology, and inheritance. Speaking English as a native language is the only variable to be significantly associated with evolutionary knowledge at Time 2.

These results are in contrast to those in Time 1 where the only significant variable associated with all biological constructs was speaking English as a native language.

Particularly, with regards to evolutionary knowledge at Time 1 and high parental levels of education.

Table 9.11. Partial correlations between parent demographic data and biological construct at Time 2

(df=74)	Biodiversity	Ecology	Inheritance	Evolution
Number of adults collapsed to plus 4	-0.252*	0.013	-0.052	0.060
Number of older children at home	-0.147	-0.118	-0.025	-0.043
Number of younger children at home	0.063	-0.141	-0.149	-0.054
Children who attended preschool yes/no	0.102	0.172	0.165	0.083
Free school meals	-0.021	-0.090	-0.210	-0.114
English native	-0.398**	-0.334**	-0.235*	-0.264*
Education level of mother	0.229*	0.182	0.463**	0.162
Education level of father	0.169	0.082	0.181	0.069
Occupation level of mother	0.267*	0.209*	0.341*	0.188
Occupation level of father	0.278*	0.262*	0.273*	0.197

*p<0.05; **p<0.001

Similar partial correlations were also conducted to see the relationship between parent demographic data and general cognitive abilities at Time 2, controlling for age (Table 9.12). These correlations reveal that once again, speaking English as a native language is associated with better receptive language (BPVS, backward digit recall, number knowledge, and expressive language. In fact performance on the expressive language task and the BPVS were both associated with parental levels of education and occupation, as might be expected. Performance on executive control, however, was only associated with fathers' levels of education and occupation, whereas number knowledge was only related to mothers' occupation level.

Table 9.12. Partial correlations between parent demographic data and general cognitive ability measures at Time 2 controlling for age

(df=74)	BPVS	Digit recall	b/w digit recall	Block recall	NKT	WCST	Stroop	Expressive Language
Number of adults	-0.035	-0.059	-0.035	0.262*	-0.143	0.313*	0.024	-0.081
Number of older children at home	-0.338*	-0.050	-0.005	-0.174	-0.142	-0.046	0.456**	-0.175
Number of younger children at home	-0.052	0.034	-0.077	-0.091	-0.115	-0.173	0.117	-0.080
Preschool attendance	0.112	0.012	0.039	0.110	0.205	0.124	-0.044	0.117
SES	-0.283*	-0.056	-0.130	0.010	-0.042	0.028	0.155	-0.253*
Native English	-0.498**	-0.175	-0.258*	0.207	-0.321*	-0.212	0.108	-0.389**
Education level of mother	0.292*	0.050	0.104	0.047	0.102	0.086	-0.076	0.354**
Education level of father	0.351*	0.038	0.261*	0.030	0.089	0.022	-0.162	0.381**
Occupation level of mother	0.344*	0.172	0.136	0.000	0.262*	0.088	-0.118	0.342*

Occupation level of father	0.419**	0.128	0.290*	-0.069	0.151	0.031	-0.196	0.400**

*p<0.05; **p<0.001

9.6 Partial correlations

Correlations controlling for age at time of testing were conducted to see how far biological constructs correlate with each other at Time 2 (Table 9.13). These analyses show that for Time 2, all biological concepts are significantly correlated with other, as they were at Time 1, also.

Table 9.13. Partial correlations between all biological constructs at Time 2, controlling for age (df=126)

	Biodiversity	Ecology	Inheritance	Evolution
Biodiversity	1	0.637**	0.395**	0.651**
Ecology		1	0.340**	0.578**
Inheritance			1	0.530**
Evolution				1

*p<0.05; **p<0.001

Partial correlations were also conducted between measures of general cognitive ability to observe the extent to which these were all related to each other (Table 9.14). These correlations illustrate that BPVS and expressive language are related to all other variables except block recall and WCST. Backwards digit recall is also associated with all other variables, which is a change from Time 1 where it was not significantly associated with block recall or WCST. Finally, number knowledge is also significantly associated with inhibition and expressive language, which again, is a change from Time 1 where it was not. Overall it seems as though block recall and WCST do not have much significant association with many of the other variables at Time 2, as was also seen as Time 1.

Table 9.14. Partial correlations between all measures of general cognitive ability, controlling for age (df=126)

	BPVS	Digit recall	B/W digit recall	Block recall	NKT	WCST	Stroop	Expressive Language
BPVS	1	0.380**	0.363**	-0.005	0.317**	0.036	-0.273*	0.707**
Digit recall		1	0.426**	0.139	0.130	0.064	-0.120	0.367**
B/w digit recall			1	0.249*	0.301**	0.180*	-0.197*	0.332**
Block recall				1	0.080	0.050	-0.104	-0.001
NKT					1	0.165	-0.186*	0.336**
WCST						1	-0.075	0.064
Stroop							1	-0.196*
Ex'sive Lang								1

*$p<0.05$; **$p<0.001$

When partial correlations are conducted between the biological constructs and the general cognitive measures, it seems as though BPVS is significantly correlated to all biological concepts and cognitive measures except block recall and WCST (Table 9.15). Digit recall, backwards digit recall, and number knowledge are significantly correlated with all biological concepts. Block recall and WCST generally do not correlate with anything, yet the stroop measure is significantly negatively correlated with all biological constructs except for evolution. Finally expressive language is significantly correlated to all measures except block recall and WCST. These findings are similar to those in Time 1 where BPVS was significantly correlated with all biological constructs. Digit recall and backwards digit recall were only significantly correlated to biodiversity and evolution at Time 1, however these are not significantly correlated with all biological measures, as is number knowledge, which at Time 1 was only associated with biodiversity. Block recall is not significantly correlated with any biological construct at Time 1 or Time 2 suggesting very little association between these variables.

Interestingly, expressive language was not measured at Time 1 but at Time 2 it seems to generally have a stronger relationship with the biological constructs in comparison to the BPVS. This may be because the biological task is more expressive rather than receptive in nature.

Table 9.15. Partial correlations between measures of general cognitive ability and biological constructs at Time 2, controlling for age

Partial Correlations Time 2					
Control Variables		Biodiversity score	Ecology score	Inheritance score	Evolution score
Age at time of testing	BPVS	0.560**	0.471**	0.306**	0.366**
	Digit Recall	0.289**	0.341**	0.237*	0.293**
	Backward digit recall	0.291**	0.188*	0.194*	0.228*
	Block recall	0.012	0.027	-0.049	0.080
	Number Knowledge score	0.335**	0.285**	0.170*	0.280**
	WCST	0.018	0.086	0.062	0.043

	Stroop	-0.346**	-0.289**	-0.127	-0.183*
	Expressive Language	0.611**	0.505**	0.428**	0.528**

**p<0.001; *p<0.05

9.7 Exploratory Factor Analyses

At the start of Chapter 8 analyses were conducted to check the validity and internal consistency of the measure testing biological knowledge in children. The results from these analyses revealed moderate alphas but because these were not that high, there is the possibility that there might be some more differentiated structure. For this reason, exploratory factor analyses were attempted using Kaiser's criterion, primarily as checks for the biological structures. Ideally, all core knowledge elements relating to all four biological constructs would have been included in the model to see whether or not a four-factor solution was found, one factor for each biological construct, thus confirming validity of the measurements. However, as there were 86 core knowledge elements in total, this would have required a much larger sample size than the current 129 participants. For this reason alternative exploratory factor analyses were conducted for each biological construct so see if any sub-structures appeared within each construct, which may give some insight into which sub-concepts or questions children performed better or worse on, and provide some insight into the moderate alpha scores described in the previous chapter.

Exploratory factor analyses were conducted for each biological construct at Time 1 and Time 2 using Kaiser's criterion. All elements related to each of the four constructs were entered in separate analyses, except elements with zero variance, which were excluded (Q11-I6sav;

Q11-I6pon; q30-Ev38pon). To clarify with an example, a factor analysis was conducted to examine biodiversity at Time 1 using all Time 1 biodiversity elements (except those with zero-variance).

The data in every instance were plainly factorable and as Bartlett's approximate chi-square was a very sensitive measure and thus less reliable, KMO statistics were used as a reliable measure of the factorability of the data. In every instance the KMO values were very high suggesting a strong relationship, however the resulting factor loadings, which often revealed seven factor solutions for each biological construct, were not sensibly interpretable. For these reasons scree plots were consulted as an alternative way to determine the optimal number of factors. These plots all suggested either a two or three factor solution for all biological constructs at Time 1 and at Time 2. Based on the analyses attempted here, confirmatory factor analyses were computed using the two or three solutions suggested by the scree plots. The results from the exploratory factor analyses can be seen in Table 9.16.

Table 9.16. Results of exploratory factor analysis for each biological construct at Time 1 and Time 2

		sample size	number of variables	KMO	Bartlett's chi-sq	Factors extracted	Factors suggested by scree plot	Cumulative variance	variance explained by first factor	overall communalities
Time 1	Bio	137	26	0.795	(325) 1274.950**	7	2	64.84%	25.88%	>0.540
	Eco	137	20	0.746	(190) 665.330**	6	2	32.67%	12.77%	>0.500
	Inh	137	12	0.658	(45) 186.400**	4	2	61.32%	27.76%	>0.450
	Evo	137	20	0.706	(171) 765.920**	7	3	67.71%	22.60%	>0.410
Time 2	Bio	129	26	0.824	(325) 2094.790**	7	2	73.79%	36.17%	>0.585
	Eco	129	20	0.672	(190) 1091.830**	7	3	71.74%	25.80%	>0.500
	Inh	129	12	0.732	(66) 427.010**	5	3	53.25%	30.34%	>0.327
	Evo	129	21	0.646	(171) 686.400**	7	3	73.83%	25.12%	>0.587

**p<0.001

9.8 Confirmatory Factor Analyses

Based on the exploratory factor analyses conducted above on each of the biological constructs at Time 1 and Time 2, it seemed appropriate, given the results of these prior analyses, to conduct confirmatory factor solutions using the number of factors specified from the scree plot from the previous analyses (Table 9.17). This seemed like a sensible strategy to see if anything interpretable could be gained from looking at this data. Also based on the earlier reliability analyses, the alpha values implied that there is some high inter-correlation between the alphas of different elements meaning that one may not actually see a separable structure, particularly given that all elements associated to a biological construct were chosen on the very basis of their association with the construct. Table 9.17 summaries the results.

Table 9.17. Confirmatory factor analyses for each biological construct at Time 1 and Time 2

		sample size	No. of variables	KMO	Bartlett's chi sq	Factor extracted	Cumulative variance	variance explained: 1st factor	variance explained: 2nd factor	variance explained: 3rd factor	overall communalities
Time 1	Bio	137	26	0.795	(325) 1274.95**	2		22.290%	33.860%		>0.021
	Eco	137	20	0.746	(190) 665.33**	2	32.670%	19.894%	12.776%		>0.109
	Inh	137	12	0.658	(45) 186.40**	2	38.923%	2210.800 %	16.815%		>0.078
	Evo	137	20	0.706	(171) 765.92**	3	43.355%	17.404%	17.205%	17.205%	>0.024
Time 2	Bio	129	26	0.824	(325) 2094.79**	2	43.000%	21.970%	21.410%		>0.202
	Eco	129	20	0.672	(190) 1091.83**	3	45.360%	16.163%	15.930%	13.260%	>0.198
	Inh	129	12	0.732	(66) 427.01**	3	53.250%	20.820%	18.888%	13.550%	>0.248
	Evo	128	21	0.758	(171) 686.93	3	48.78%	23.000%	13.60%	12.32%	>0.162

**p>0.001

The confirmatory factor analyses were conducted on the basis that there may be an identifiable sub-structure within each of the biological constructs, which in turn may have had an effect on the way that children scored on the various elements for each construct. However, the results of these analyses are not clear cut enough to warrant further exploration of the data, as no clear factor structure was seen, and the majority of the variance loaded onto the first factor. Communalities for many of the variables were also very low implying this method of analysis may not be appropriate. The variance predicted by a 2-factor solution was not always high, suggesting that these factors may not reflect a subgroup of children's knowledge or way of thinking, rather it may simply be an artefact of the data.

Regardless, the alphas from Chapter 8 suggest a high degree of within scale coherence. Given this fact, it was clear that the elements within constructs as well as the constructs themselves are highly related suggesting the elements related to a construct fall into one factor – that of each biological construct. Hence observing a clear factor structure would have been somewhat unlikely.

9.9 Interim summary & discussion of results: Time 1 & Time 2

Thus far, analyses have only been viewed cross-sectionally in an effort to try and map out the route of conceptual change and progression. There have been various useful results as a consequence of this, one of the main results being about the relative strength of knowledge children hold on biodiversity concepts, regardless of the context in which these concepts are held. Both at Time 1 and at Time 2, children across all cohorts perform significantly better on biodiversity than any other biological concept. This may be due to the types of core knowledge structures that are implicated in biodiversity concepts. Many of the ideas relate to taxonomy which build on the need to categorise, and as described in Chapter 4, there is a wealth of research suggesting that children are proficient in categorising from a very young age, which would explain their strengths in biodiversity over other biological constructs. It may also be that the labelling of organisms that is required prior to any classification could explain why language is also significantly correlated to biodiversity, as is age, which seems to serve as a proxy for the amount of experience a child has.

Figure 9.1 illustrates how children's performance across cohorts is very similar for ecological and evolutionary constructs. These two constructs are unique in that the core knowledge structures focus on global thinking for example, thinking about an ecosystem, or the process of natural selection. These types of ideas are also dependent on children's ability to gauge *change over time* and would require the underlying domain general ability of maintaining multiple ideas (Hipkins et al., 2008; Vosniadou, 2014) which would explain why children perform similarly in these areas. The relative jump seen at Time 2 where children's knowledge of evolution generally overtakes their knowledge of ecology may also be a result of the fact that at Time 2, the new NC (DfE, 2014) was introduced and teachers may have been emphasising these new evolutionary concepts to children in science lessons relative to ecological ones. These two concepts also focus less on categorisation ability but more on processes or causal mechanisms behind biological systems and phenomena. As such, children might not be as proficient in these types of concepts relative to biodiversity.

In contrast, children's knowledge of inheritance was significantly less than any other biological construct, with no significant change in this knowledge across time points. Given past research suggesting children have relatively coherent understanding of inheritance (e.g. Springer, 1999; Carey, 1985) it was expected that this would be an area that children would perform relatively well in. However, as this study took a broader and more robust approach in looking at inheritance concepts, it became clear that children actually have very limited knowledge. This may reflect that at primary school level, children are not formally taught sexual reproduction and so children can only reflect on ideas of inheritance up to a point. This may also explain the lack of any significant change or improvement across time points or indeed across cohorts. The very few concepts children do seem to understand with regards to inheritance might be those around population, hence the high correlations with ecology and evolution.

Results in Figure 9.1 also highlight a key issue with regards to the longitudinal effects of testing children. The pattern of results show consistent improvement of children's knowledge of biological constructs (aside from inheritance) from Time 1 to Time 2 where for instance, children at Time 2 in Cohort 1 are performing to a similar level to children in Cohort 2 at Time 2. This consistent linear increase suggests that testing alone has improved children's knowledge of these biological constructs. The testing procedures were interview based, getting children to think about organisms and processed of various biological phenomena and to explain their understanding. This would suggest that when children are merely given the opportunity to reflect upon reasons for their ideas more deeply, and engage in dialogue about these ideas, conceptual development and progression improves (discussed in Chapter 11).

Correlational analyses revealed outcomes that were also very informative about the types of variables from the general cognitive measures that were significantly associated with biological constructs. These suggested verbal measures including receptive and expressive vocabulary, and short-term verbal working memory were key, implying language may be a key factor in the process of conceptual change. Similarly executive control and number knowledge were also influential at Time 2. However, across the two time points, block recall and cognitive flexibility were not significantly associated, and semantic inhibitory control also had no significant association at Time 1 (although more at Time 2). Thus in terms of moving forward and trying to *explain* the changes in biological knowledge described earlier, it becomes a useful exercise to try to predict the amount of variance explained by general cognitive ability measures for each biological construct at Time 1 and Time 2. The correlational outcomes provided a good way to triage the variables that could be used in regression analyses reported in Chapter 10.

CHAPTER 10: MODELLING ANALYSES

10.1 Overview

The relationships in the data that became evident after correlational analyses from Chapter 8 and Chapter 9 were analysed in more detail. It was decided that hierarchical regression analyses with a single dependant variable would provide a useful avenue to investigate the impact of general cognitive influences as a first step, then the effect on the resulting models after including other aspects of biological understanding. This made it possible to check the relative impact of general and specific predictors, and in part, how far the impact of general cognitive predictors was mediated through more specific influences. These analyses used data from the whole sample, as far as data were available, so that there were no issues of infringing conventions on the number of cases per predictor used in the regression models.

Two-stage hierarchical regression models were computed using each of the four biological constructs (biodiversity, ecology, inheritance, and evolution) as dependent variables separately, and using the general cognitive ability measures alongside chronological age as predictors in the first stage of the model. The remaining biological constructs were then hierarchically added as predictors in the second stage of the model. This includes a total of 8 predictors at stage one, and 11 predictors at stage two. The rationale behind these models was to try and explain the changes and developmental patterns across the two time points observed in the previous chapter. Note that age was included as a predictor in the models and was a constant feature in these analyses because it removed the variance attributable to participants being in different cohorts.

This chapter outlines the first preliminary regression models conducted using all general cognitive predictors at step one, and the remaining biological constructs as predictors in step two for each biological construct. The results from these analyses highlighted the significant predictors in each model, as well as highlighting the extraneous variables which seemingly have very little influence on the outcome variable. These, taken together with previous results from the earlier correlational analyses, led to later regression analyses using only predictor variables that consistently had a significant effect on the outcome variables, both at Time 1 and at Time 2.

Regression models predicting the change at Time 2 from Time 1 variables were then attempted for each biological construct using the significant predictors that emerged from the earlier models. Also, due to the fact that earlier correlational analyses suggested that demographic variables were significantly associated with biological knowledge at Time 2, preliminary hierarchical regression models using demographic variables in step one and remaining biological predictors at step two were computed. These models highlighted the significant demographic predictors of biological constructs at Time 1 and Time 2.

As with the cognitive predictors in the earlier regression models, the extraneous demographic variables that did not significantly contribute towards the variance in the biological outcome variables were removed from the subsequent analyses. This now meant that significant cognitive predictors and significant demographic predictors derived from preliminary analyses could be included in composite hierarchical regression models for each biological construct at Time 1 and Time 2, to see which out these significantly associated indictors significantly predicted the variance of each biological construct (in step one) and how this influence changed when remaining biological constructs were subsequently included in the model (at step two).

10.2 Preliminary regression analyses

10.2.1 Preliminary regression results from Time 1

The variables that were included in the first stage of the hierarchical regression for each biological construct were: age, BPVS, digit recall, backwards digit recall, block recall, NKT, WCST, and Stroop task scores. The second stage of the model included the composite mean scores from the remaining biological constructs to the one being predicted.

Findings revealed that all biological constructs produced strong and significant regression models at Time 1 ($AdjR^2$ >0.6 in all cases for the final model). The preliminary models for biodiversity (Table 10.1), ecology (Table 10.2), inheritance (Table 10.3), and evolution (Table 10.4) at Time 1 can be viewed below:

Generally, when only general cognitive predictors and age in months are included in the models, BPVS is the strongest predictor (aside from biodiversity where age is also a significant predictor). When other biological concepts are included in the model, BPVS drops out as a predictor in all of the models to be replaced by some of the other biological concepts. For biodiversity at Time 1, age and backwards digit recall also remain significant predictors, whilst for evolution at Time 1 digit recall remains a strong predictor in model 2.

Table 10.1. Preliminary hierarchical regression model for biodiversity at Time 1

Biodiversity Time 1		
Model 1: AdjR^2=0.644 (p<0.001)		
	Beta	sig
Age	0.325	0.381
BPVS	0.288	**0.006**
Digit recall	0.044	0.548
b/w digit recall	0.188	**0.035**
Block recall	0.046	0.536
NKT	0.026	0.706
WCST	-0.031	0.581
Stroop	0.035	0.575
Model 2: AdjR^2=0.773 (p<0.001)		
Age	0.274	**0.002**
BPVS	0.027	0.766
Digit recall	0.026	0.671
b/w digit recall	0.156	**0.030**
Block recall	-0.003	0.958
NKT	0.024	0.655
WCST	-0.062	0.177
Stroop	0.038	0.444
Ecology	0.333	**<0.001**
Inheritance	-0.015	0.799
Evolution	0.228	**0.001**

Table 10.2. Preliminary hierarchical regression model for ecology at Time 1

Ecology at Time 1		
Model 1: AdjR^2=0.397 (p<0.001)		
	Beta	sig
Age	0.077	0.583

BPVS	0.560	**<0.001**
Digit recall	-0.054	0.579
b/w digit recall	0.017	0.886
Block recall	0.104	0.281
NKT	0.005	0.957
WCST	0.041	0.579
Stroop	0.002	0.983
Model 2: AdjR^2=0.662 (p<0.001)		
Age	-0.173	0.116
BPVS	0.201	0.063
Digit recall	-0.098	0.184
b/w digit recall	-0.067	0.447
Block recall	0.101	0.166
NKT	-0.016	0.810
WCST	0.060	0.286
Stroop	-0.015	0.805
Biodiversity	0.496	**<0.001**
Inheritance	0.266	**<0.001**
Evolution	0.216	**0.011**

Table 10.3. Preliminary hierarchical regression model for Inheritance at Time 1

Inheritance at Time 1		
Model 1: AdjR^2=0.293 (p<0.001)		
	Beta	sig
Age	0.231	0.131
BPVS	0.519	**0.001**
Digit recall	-0.044	0.670
b/w digit recall	-0.124	0.320
Block recall	-0.120	0.252
NKT	0.029	0.760
WCST	-0.070	0.382

Stroop	0.010	0.907

Model 2: AdjR^2 = 0.455 (*ps* <0.001)		
Age	0.183	0.191
BPVS	0.215	0.119
Digit recall	-0.052	0.577
b/w digit recall	-0.147	0.192
Block recall	-0.174	0.060
NKT	0.028	0.740
WCST	-0.103	0.148
Stroop	0.014	0.857
Biodiversity	-0.035	0.799
Ecology	0.430	**<0.001**
Evolution	0.205	0.057

Table 10.4. Preliminary hierarchical regression model for evolution at Time 1

Evolution Time 1		
Model 1: AdjR^2=0.454 (*p*<0.001)		
	Beta	sig
Age	0.127	0.342
BPVS	0.361	**0.006**
Digit recall	0.158	0.087
b/w digit recall	0.109	0.321
Block recall	0.055	0.550
NKT	0.001	0.992
WCST	0.071	0.315
Stroop	-0.016	0.839
Model 2: AdjR^2=0.626 (*p*<0.001)		
Age	-0.045	0.696
BPVS	0.047	0.683
Digit recall	0.160	**0.038**

b/w digit recall	0.052	0.579
Block recall	0.030	0.701
NKT	-0.014	0.841
WCST	0.083	0.163
Stroop	-0.031	0.632
Biodiversity	0.375	**0.001**
Ecology	0.239	**0.011**
Inheritance	0.141	0.570

The results from Tables 10.1–10.4 highlighted some key pieces of information. Firstly, receptive language seems to be a very strong predictor and its influence was significant for all models at step 1. This is even despite the fact that the BPVS task controls for age. Secondly, although receptive language is a significant predictor at step 1 for all models, in step 2 when other constructs of biological understanding are included, the BPVS drops out as a predictor in all step 2 models. This suggests that language ability *per se* is not important; rather it is its specific manifestation in different forms of biological knowledge. If one considers this notion in further detail, it becomes relatively clear that the results are not simply attributable to the method of assessment because if this were true then all constructs of biological knowledge would carry approximate equal predictive weight. However, as this is not apparent, it would suggest that the language effect seen here is more specific.

Additionally, receptive language is most strongly predictive of ecology. Upon further inspection, ecology is the only biological construct that is predictive of all other biological constructs, and is associated with each of them in turn. This would seem to suggest that language acquired around ecological concepts might somehow be central, and perhaps be the first step towards driving conceptual progression across biological concepts.

In sum, it seems to be that ecology is the main driver for other areas of understanding, with language acquired here then impacting on biodiversity and inheritance. This is also supported by the pattern of relationships seen with receptive language. Biodiversity is significantly related to age in the step 2 model suggesting an effect of experience not evident in the other aspects of knowledge, and to executive control. Interestingly, in Chapter 8 it was observed that at Time 1, biodiversity and ecology are the two areas on which children perform the best in general. The fact that biodiversity is more strongly predictive of ecology than vice versa in these models (see Figure 10.1), may indicate that growth in understanding of biodiversity forms a basic stratum on which ecology then builds.

Inheritance produces the weakest model, although this model is still significant at $p<0.001$. This finding reflects the results from the previous chapters where children have seemingly little knowledge in this area in comparison to other areas of biology.

Finally, it seems as though general cognitive abilities have ostensibly little association with conceptual change in biological understanding; aside from executive control in the biodiversity model and short-term verbal memory in evolution. Instead it could be argued that both ecology and biodiversity are the main driving factors with language developing at ecology passing through to biodiversity, and then biodiversity and ecology might encourage the development of evolution and inheritance concepts. Regardless, the main finding to arise from these analyses is that language plays a significant part in the conceptual development and progression of biological knowledge. The main points from these analyses shown in Table 10.1-10.4 are captured in Figure 10.1 below.

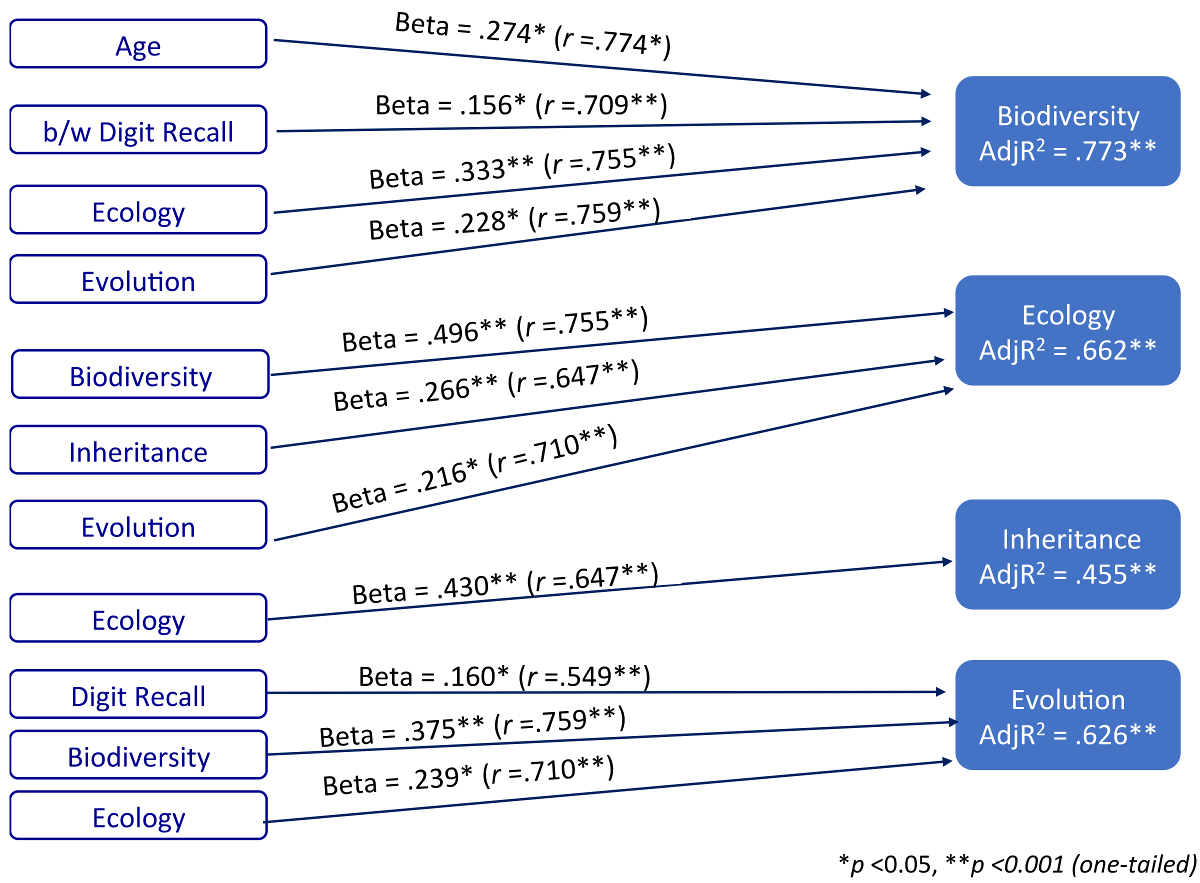


Figure 10.1. Preliminary hierarchical regression models for biological constructs at Time 1

10.2.2 Preliminary regression results from Time 2

The same two-stage hierarchical multiple regressions were conducted for scores for each biological construct within Time 2 data. The same seven general cognitive variables and age were used as variables for the first step, with the inclusion of the remaining biological variables at step two. Note that for the Time 2 models expressive language scores were also included as a general cognitive predictor at step 1, hence there were 9 predictors in the first stage, and 12 predictors in the second stage.

Findings revealed that significant regression models were produced for all of the biological concepts at Time 2 ($AdjR^2$ >0.6 for most cases for the final model, and $AdjR^2$ >0.4 for the final inheritance model). The strongest model was for biodiversity (Table 10.5), and the weakest model was for inheritance (Table 10.7):

Table 10.5. Preliminary hierarchical regression model for biodiversity at Time 2

Biodiversity Time 2		
Model 1: $AdjR^2$=0.724 (p<0.001)		
	Beta	sig
Age	0.067	0.476
BPVS	0.190	0.101
Digit recall	0.022	0.716
b/w digit recall	0.037	0.602
Block recall	-0.024	0.659
NKT	0.109	0.148
WCST	-0.043	0.370
Stroop	-0.112	**0.032**
Expressive language	0.474	**<0.001**
Model 2: $AdjR^2$=0.797 (p<0.001)		
Age	0.039	0.628
BPVS	0.092	0.358
Digit recall	-0.034	0.509
b/w digit recall	0.069	0.271
Block recall	-0.039	0.411
NKT	0.049	0.456
WCST	-0.045	0.279
Stroop	-0.086	0.059
Expressive language	0.249	**0.025**

	Ecology	0.250	**0.001**
	Inheritance	-0.007	0.894
	Evolution	0.280	**<0.001**

Table 10.6. Preliminary hierarchical regression model for ecology at Time 2

	Ecology Time 2		
Model 1: $AdjR^2$=0.627 (p<0.001)			
		Beta	sig
	Age	0.141	0.198
	BPVS	0.180	0.179
	Digit recall	0.146	**0.032**
	b/w digit recall	-0.099	0.238
	Block recall	-0.003	0.965
	NKT	0.116	0.187
	WCST	0.023	0.680
	Stroop	-0.102	0.092
	Expressive language	0.368	**0.009**
Model 2: $AdjR^2$=0.702 (p<0.001)			
	Age	0.122	0.213
	BPVS	0.068	0.575
	Digit recall	0.123	**0.050**
	b/w digit recall	-0.107	0.155
	Block recall	-0.007	0.901
	NKT	0.050	0.523
	WCST	0.042	0.400
	Stroop	-0.060	0.273
	Expressive language	0.091	0.502
	Biodiversity	0.366	**0.001**
	Inheritance	-0.014	0.833
	Evolution	0.226	**0.015**

Table 10.7. Preliminary hierarchical regression model for inheritance at Time 2

Inheritance Time 2		
Model 1: AdjR^2=0.322 (p<0.001)		
	Beta	sig
Age	-0.062	0.675
BPVS	-0.056	0.757
Digit recall	0.085	0.364
b/w digit recall	0.057	0.615
Block recall	-0.079	0.357
NKT	0.035	0.766
WCST	0.013	0.866
Stroop	-0.017	0.835
Expressive language	0.613	**0.001**
Model 2: AdjR^2=0.410 (p<0.001)		
Age	-0.041	0.765
BPVS	-0.141	0.410
Digit recall	0.053	0.552
b/w digit recall	0.066	0.537
Block recall	-0.107	0.186
NKT	-0.017	0.878
WCST	0.020	0.782
Stroop	-0.021	0.792
Expressive language	0.384	**0.043**
Biodiversity	-0.021	0.894
Ecology	-0.028	0.833
Evolution	0.507	**<0.001**

Table 10.8. Preliminary hierarchical regression model for evolution at Time 2

Evolution Time 2		
Model 1: AdjR^2=0.571 (p<0.001)		
	Beta	sig
Age	-0.030	0.800
BPVS	0.185	0.197
Digit recall	0.072	0.332
b/w digit recall	-0.022	0.809
Block recall	0.054	0.431
NKT	0.114	0.226
WCST	-0.014	0.814
Stroop	-0.003	0.964
Expressive language	0.492	**0.001**
Model 2: AdjR^2=0.712 (p<0.001)		
Age	-0.072	0.459
BPVS	0.084	0.481
Digit recall	0.010	0.866
b/w digit recall	-0.029	0.698
Block recall	0.084	0.139
NKT	0.036	0.641
WCST	-0.005	0.917
Stroop	0.068	0.210
Expressive language	0.071	0.596
Biodiversity	0.398	**<0.001**
Ecology	0.219	**0.015**
Inheritance	0.248	**<0.001**

Based on the regression models from Time 2, the idea that specific language is influential in conceptual developmental for biological phenomena still remains. With the inclusion of expressive language in the Time 2 models, BPVS is never a significant predictor for any of the models like it was at Time 1. This would suggest that it is plainly the encoding of ideas, not simply recognition of language that is important in conceptual development. The importance of this is further signalled by the fact that unlike BPVS at Time 1, expressive language does not always drop out of the step 2 models (although its influence is always weaker at step 2), which is consistent with the premise that specific biological language might be key predictor of biological knowledge.

The exact relationship between the biological constructs in the step 2 models is different to Time 1, possibly in part because of the variance explained by the newly included expressive language measure, but perhaps also because the drivers and their effects themselves have to some extent shifted. Most markedly, evolution has overtaken ecology as the strongest area of performance after biodiversity, and has now displaced ecology as central in the sense of predicting all other biological constructs, and also being predicted by them.

The ecology model seems to be the most distinct as when the biological concepts are included in the model, expressive language drops out and only digit recall, evolution, and biodiversity remain as significant predictors. The pattern for evolution concepts is also unique, expressive language drops out of the first model and the only significant predictors are the remaining biological concepts. If the Time 1 data are taken to be accurate, the implication may be that ecology, having built on biodiversity and fed into evolution, has now led to more pronounced growth in evolution. Inheritance is still relatively speaking an outlier, again producing the weakest model, but the influence of evolution and expressive language on inheritance would indicate that this should be the next area of growth. Also, the shift from the significant effect of ecology at Time 1, to evolution at Time 2 would suggest there might be some type of reciprocal relationship between ecology and evolution.

With regards to the general cognitive measures, none of those that were observed at Time 1 survive to remain intact to Time 2. Instead, one other (digit recall) has come in. This may indicate that the influences of general cognitive abilities are all weak chance effects, but perhaps more likely that they have marginal influences that are hard to capture accurately, because the data are noisy to some extent. Figure 10.2 below summarises the final models for each biological construct.

To summarise, when only general cognitive predictors and age are included in the step 1 models, expressive language seems to be a core predictor. Receptive language is not significantly predicting any variance as it did at Time 1 in any model. When other biological concepts are included in the model, expressive language does not drop out of all of the models, but remains as a significant predictor of biodiversity and inheritance.

Contrary to Time 1 where it seemed that ecology was the main driver for other areas of understanding, with language acquired here then impacting on biodiversity and evolution, the results at Time 2 reveal that evolution is the main driver of change. Expressive language feeds into biodiversity and inheritance, and then biodiversity seems to be feeding into ecology specifically, whilst inheritance feeds into evolution concepts along with biodiversity (note that age was a significant factor for the biodiversity model at Time 1).

There is definitely a shift in language from ecology at Time 1, to biodiversity at Time 2. The fact that expressive language is also a significant predictor of inheritance is unsurprising on the account that reproduction is not formally taught in the primary school curriculum and so children are likely to try to articulate themselves as best they can without knowing anything much about the processes involved; those that were able to do this best are likely to have high expressive language.

It seems as though general cognitive abilities have very little association with conceptual change in biological understanding; instead it may be biologically specific language that drives forward change, in the second time of testing, this may be children's expressive language specifically.

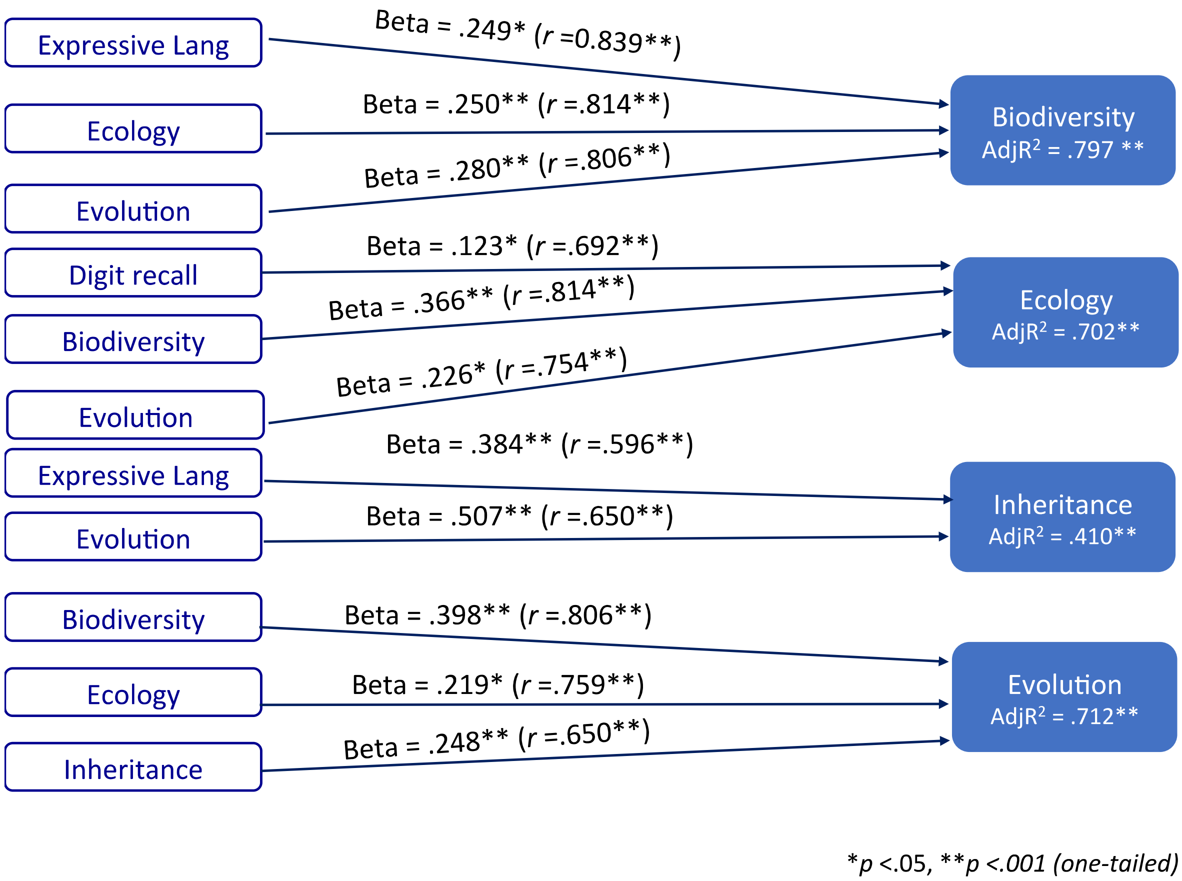


Figure 10.2. Summary of the final regression models for all biological constructs at Time 2

10.3 Analyses post-preliminary results

The preliminary regressions revealed that collinearity was a potential issue as many of the variables had collinearity values of >0.7. This suggested that there was a high degree of correlation between some of the variables included in the model. To account for this, and also the fact that many variables included in the analysis did not have a significant contribution to the models, and would therefore potentially be adding noise to the final results, a possible solution was to exclude any variable from the analysis that consistently did not contribute a significant amount of variance to any of the models across Time 1 or

Time 2, and could therefore be classed as an extraneous variable. These extraneous variables were: block recall, NKT, and WCST measures.

10.3.1. Time 1

The two-stage hierarchical multiple regression was conducted for each biological concept in turn using age, BPVS, digit recall, backwards digit recall, and Stroop task measures as dependent factors in the first stage and including the remaining biological concepts as predictors in the second stage.

As with the preliminary analyses, significant regression models were produced for all of the biological concepts (AdjR^2 >0.62 for the majority of models, and AdjR^2 >0.44 for inheritance). For biodiversity, the results were the same as the preliminary model with age, backwards digit recall, ecology and evolution being significant predictors at step 2 (AdjR^2=0.775, p<0.001), after BPVS dropped out at step 1. Thus this model has shown a slight increase in the variance predicted by 2%. For ecology, the results were again the same as those at the preliminary stage with BPVS dropping out at step 1 to reveal the remaining biological constructs as the only significant predictors at step 2 (the variance predicted by this model was equal to that in the preliminary model: AdjR^2=0.662, p<0.001), and also for inheritance where BPVS dropped out at step 1 to reveal ecology as the only significant predictor at step 2 (a decrease of 12% in the variance predicted by this model: AdjR^2=0.442, p<0.001). For the evolution model however, there was a change from the preliminary models (Table 10.4). At step 1 only BPVS remains as significant predictor, however at step 2 this drops out of the model and only biodiversity and ecology are the significant predictors (AdjR^2=0.629,

p<0.001). This is different from the results of the Time 1 preliminary model where digit recall remains a significant predictor at step 2 and 62.6% of the variance was explained; in this model there was an increase by 2% implying the removal of nuisance variables has strengthened the model slightly.

Figure 10.3 below illustrates the final regression models for all four biological constructs at Time 1 after the removal of nuisance variables.

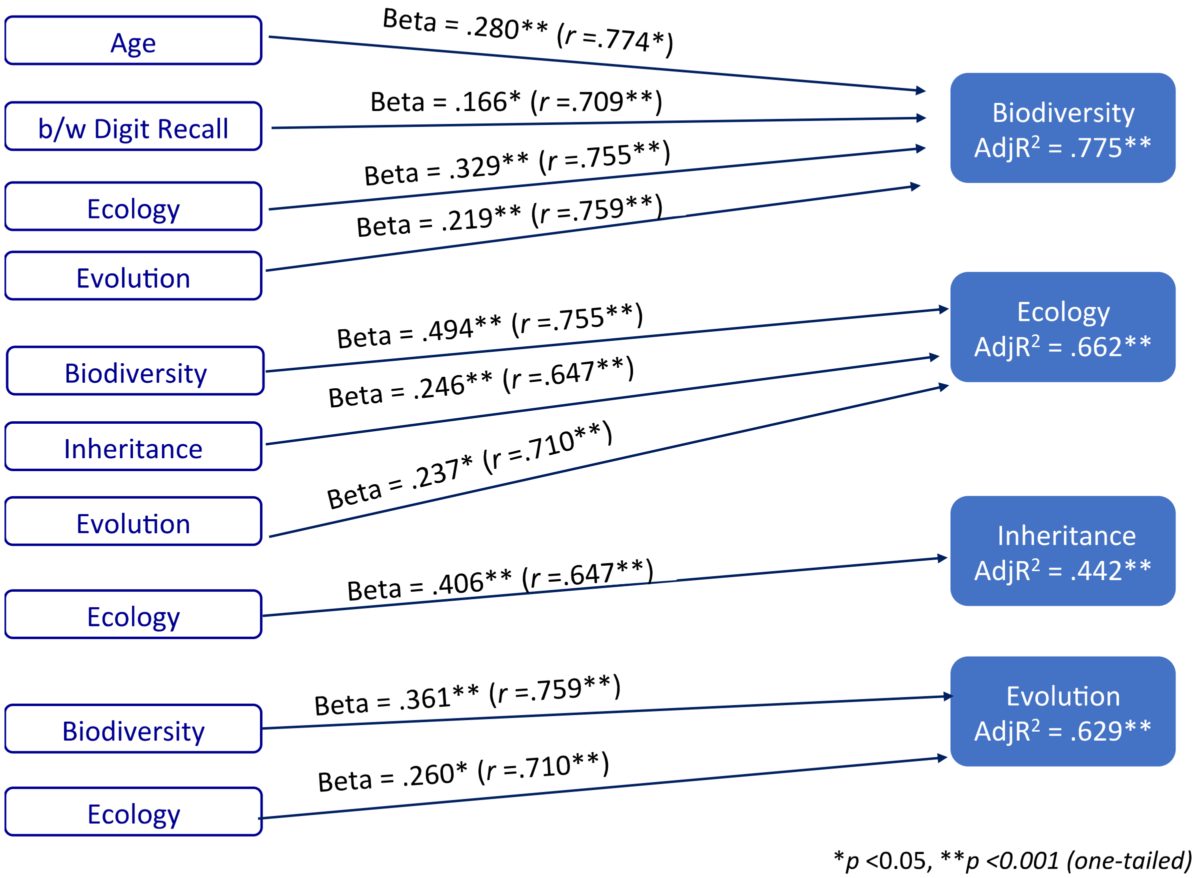


Figure 10.3. Final regression models for biological constructs at Time 1 after nuisance variables were removed

10.3.2. Time 2

The same analyses were run for Time 2, with the removal of the extraneous variables: block recall, NKT, and WCST measures.

Findings revealed that significant regression models were produced for all of the biological concepts ($AdjR^2$>0.69 in the majority of cases and $AdjR^2$>0.42 for inheritance). For biodiversity, the removal of the extraneous measures led to BPVS now becoming a significant predictor at step 1 alongside Stroop and expressive language. However at step 2, BPVS drops out and only expressive language remains with ecology and evolution as significant predictors for step 2 just like in the preliminary model (a slight decrease in the variance predicted by 10%: $AdjR^2$=0.795, p<0.001). Note that at step 2 the Stroop measure is borderline significant (beta=-0.086, p=0.052).

For the ecology model, digit recall is no longer a significant predictor at step 1 or 2 as it was in the preliminary model. Instead at step 1, expressive language is the only significant predictor, but drops out at step 2 leaving only biodiversity and evolution as predicting the majority of the variance: $AdjR^2$=0.699, p<0.001. This is a slight decrease of 5% from the amount of variance predicted in the preliminary model for ecology at Time 2.

The results for inheritance are the same as those of the preliminary model at step 1 with expressive language as the only significant predictor, however unlike the preliminary model, expressive language drops out at step 2 to leave evolution as the only remaining significant predictor of the model: $AdjR^2$=0.420, p<0.001), which is an increase of 10% of the variance explained from the preliminary model for inheritance at Time 2.

Lastly, the removal of nuisance variables for the evolution model led to the same findings as the preliminary model, whereby expressive language as the only significant predictor at step 1, drops out at step 2 to reveal the remaining biological constructs as significant predictors. The amount of variance for this model remains exactly the same as it did for the preliminary model ($AdjR^2$=0.712, p<0.001). Figure 10.4 illustrates the final regression models for all biological constructs at Time 2 with the exclusion of nuisance variables.

These final models revealed similar findings to those in the preliminary models except for the fact that expressive language is no longer a significant predictor of inheritance at Time 2, and digit span is no longer a significant predictor of ecology at Time 2. Overall however, the regression analyses warrant the same interpretation of those models in the preliminary analyses. The expressive language measure no longer being a significant predictor of inheritance at Time 2 may reflect that in the preliminary models, this was a marginal effect given the relatively large number of predictors in relation to the sample size. Figure 10.4 summarises these findings.

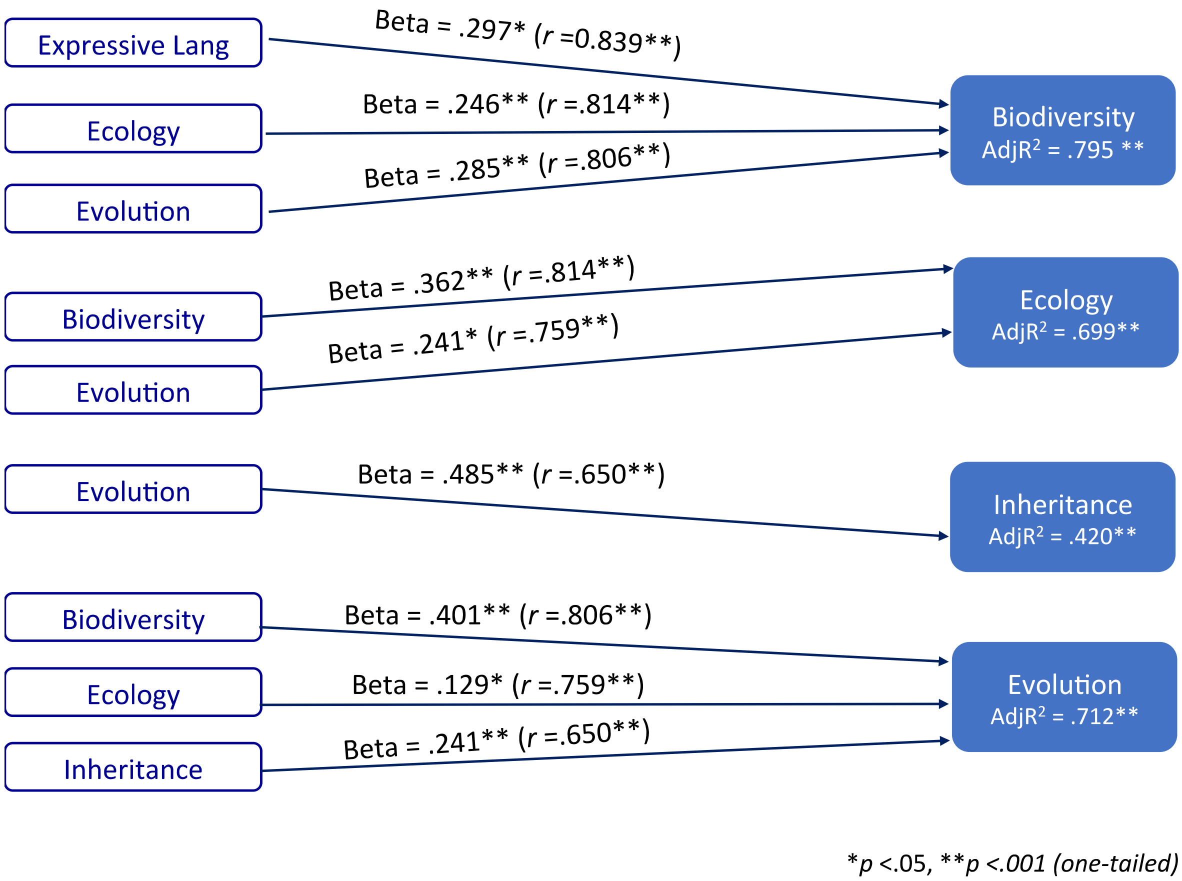


Figure 10.4. Final hierarchical regression models for biological constructs at Time 2 with nuisance variables removed

However, one issue with the models at Time 2 was the fact that expressive language was included in the model, whereas for Time 1 it was not. In order to make the models at Time 2 more comparable to those in Time 1, the same hierarchical regression was conducted again for Time 2, this time with the exclusion of expressive language, to see how similar the models may have looked. This exploratory analysis revealed that the same significant predictors remained at Time 2 except expressive language was now replaced by BPVS (see Figure 10.5 for the final regression models).

These models suggest that expressive language and receptive language have similar effects and therefore highlight the importance of specific language. Also as described earlier, they illustrate that it is the specific manifestation of language ability in different aspects of biological knowledge that is important, rather than language ability per se. This is evidenced by the fact that unlike BPVS, expressive language does not drop out at the Time 2 models suggesting that it is the encoding of ideas, and not simply the recognition of language, that is important for conceptual development.

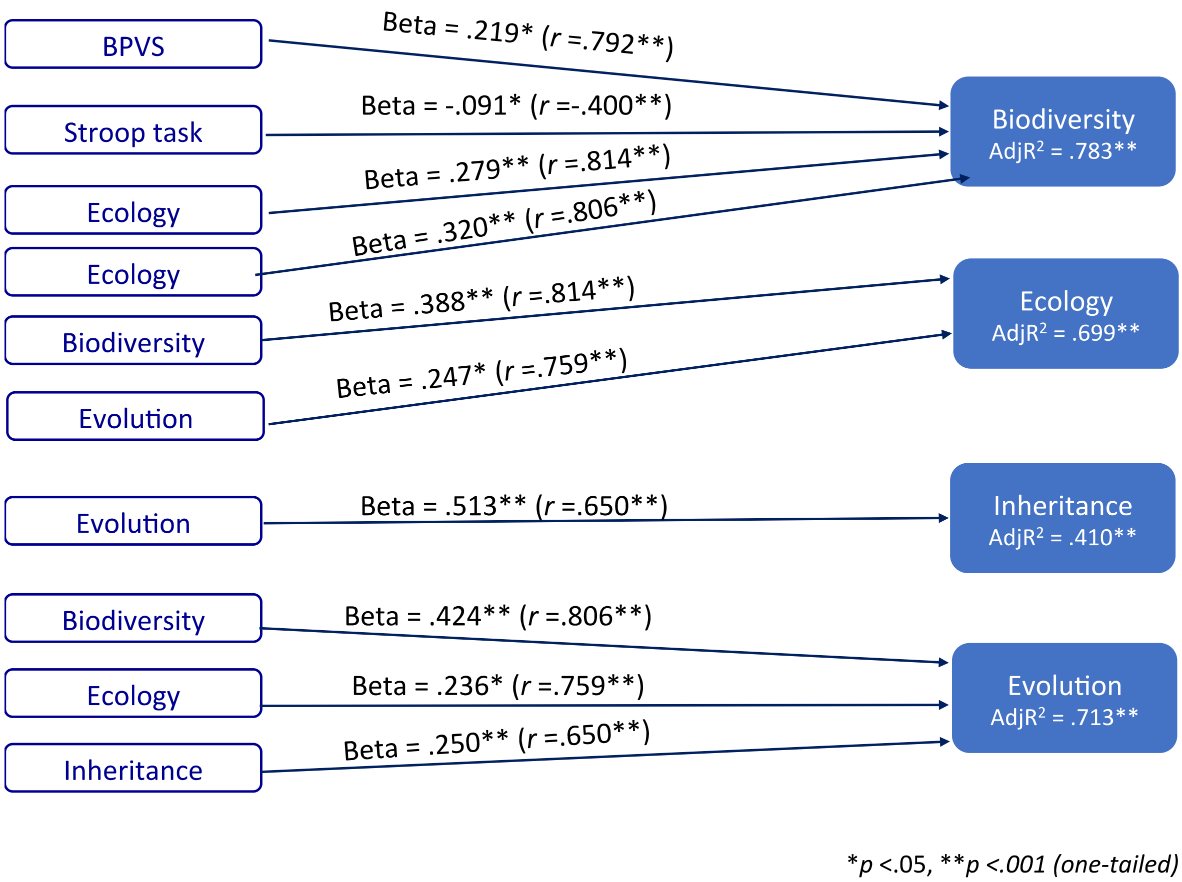


Figure 10.5. Hierarchical regression models at Time 2 excluding expressive language measure

10.4 Predicting change

10.4.1 Time 1 to Time 2 change

The earlier models have so far looked at predicting variance of the biological constructs at one time point. However in order to comment on a development and progression of these concepts, it became necessary to examine how far Time 1 general cognitive abilities (excluding those that were rendered extraneous during preliminary analyses) might be predictive of Time 2 biological knowledge. These analyses all produced significant models, which are shown in Tables 10.9-10.12 below.

Table 10.9. Final regression results for predicting biodiversity at Time 2

	Predicting Biodiversity Time 2		
Model 1: AdjR^2=0.652 (p<0.001)			
		Beta	sig
	Age	0.137	0.173
	BPVS Time 1	0.457	**<0.001**
	Digit recall Time 1	0.095	0.192
	b/w digit recall Time 1	0.159	0.065
	Stroop Time 1	-0.090	0.146
Model 2: AdjR^2=0.713 (p<0.001)			
	Age	-0.015	0.878
	BPVS Time 1	0.288	**0.007**
	Digit recall Time 1	0.060	0.376
	b/w digit recall Time 1	0.093	0.253
	Stroop Time 1	-0.102	0.072
	Biodiversity Time 1	0.415	**<0.001**
	Ecology Time 1	-0.067	0.438

Inheritance Time 1	0.127	0.057
Evolution Time 1	0.016	0.845

Table 10.10. Final regression results for predicting ecology at Time 2

Predicting Ecology Time 2		
Model 1: AdjR^2=0.573 (p<0.001)		
	Beta	sig
Age	0.110	0.321
BPVS Time 1	0.482	**<0.001**
Digit recall Time 1	0.144	0.073
b/w digit recall Time 1	0.127	0.183
Stroop Time 1	0.027	0.697
Model 2: AdjR^2=0.713 (p<0.001)		
Age	0.010	0.931
BPVS Time 1	0.316	**0.011**
Digit recall Time 1	0.132	0.095
b/w digit recall Time 1	0.085	0.367
Stroop Time 1	0.023	0.729
Biodiversity Time 1	0.231	0.060
Ecology Time 1	0.113	0.261
Inheritance Time 1	0.074	0.341
Evolution Time 1	-0.015	0.871

Table 10.11. Final regression results for predicting Inheritance at Time 2

Predicting Inheritance Time 2		
Model 1: AdjR^2=0.289 (p<0.001)		
	Beta	sig
Age	-0.029	0.839
BPVS Time 1	0.329	**0.032**
Digit recall Time 1	0.237	**0.023**
b/w digit recall Time 1	0.057	0.640

Stroop Time 1	-0.054	0.541

Model 2: AdjR^2=0.294 (*p*<0.001)		
Age	-0.076	0.612
BPVS Time 1	0.193	0.242
Digit recall Time 1	0.220	**0.039**
b/w digit recall Time 1	0.040	0.751
Stroop Time 1	-0.051	0.558
Biodiversity Time 1	0.037	0.819
Ecology Time 1	0.096	0.475
Inheritance Time 1	0.056	0.592
Evolution Time 1	0.109	0.390

Table 10.12. Final regression results for predicting evolution at Time 2

Predicting Evolution Time 2		
Model 1: AdjR^2=0.289 (*p*<0.001)		
	Beta	sig
Age	0.062	0.598
BPVS Time 1	0.429	**0.001**
Digit recall Time 1	0.149	0.083
b/w digit recall Time 1	0.164	0.106
Stroop Time 1	-0.017	0.810

Model 2: AdjR^2=0.294 (*p*<0.001)		
Age	-0.036	0.761
BPVS Time 1	0.253	0.053
Digit recall Time 1	0.145	0.086
b/w digit recall Time 1	0.124	0.218
Stroop Time 1	-0.020	0.772
Biodiversity Time 1	0.220	0.091
Ecology Time 1	0.078	0.097

Inheritance Time 1	0.063	0.445
Evolution Time 1	-0.053	0.596

The key findings to arise from the Time 1 to Time 2 analyses are that language seems to be at the core of development, particularly with its effect on biodiversity. BPVS at Time 1 is significantly predictive of biodiversity, ecology, inheritance, and evolution. Table 10.10 suggests that biodiversity measured at Time 1 might be marginally significant in predicting ecology at Time 2. Similarly, in Table 10.12 it seems as though biodiversity at Time 1 and ecology at Time 2 might also have a marginal influence in predicting evolution at Time 2. These findings are consistent with the earlier account that BPVS feeds into biodiversity at Time 1, which in turn feeds into ecology and evolution. However, Tables 10.9-10.12 show relatively weak effects. This might be due to the fact that Time 1 data were capturing live effects at that time, which may to some extent have played out by Time 2, resulting in weaker influences.

Inheritance is by and large an outlier however Table 10.9 shows the marginal influence of inheritance at Time 1 on biodiversity at Time 2. It may be that the everyday observations children have which have led to their ideas around biodiversity do to some extent overlap with their rudimentary ideas about inheritance, for example ideas around variation within a species. This marginal effect may be a result of such overlapping ideas that were not necessarily borne out or sufficient to gain higher inheritance scores, but still have an underlying albeit weak effect on related biological concepts such as those around biodiversity.

For the inheritance model, the only significant predictor is digit recall, as BPVS drops out at step 2. This seems somewhat interesting. Due to the lack of carry-over from Time 1 to Time 2, the relationships at Time 2 are seemingly more important which means there may be very low mean scores for inheritance at Time 1. It could be that the there is a weak influence from the other biological constructs but enough of an influence to absorb the variance that BPVS would have explained at Time 2 i.e. the combined reduction in shared variance could be enough to tip BPVS over to non-significant. This again fits with the idea that the other biological constructs all feed into inheritance concepts, but this area still remains very low throughout.

Finally for evolution, there are no significant predictors for the step 2 model, although BPVS remains marginally significant. This may be because there is quite high collinearity for the Stroop measure (0.730) suggesting that this measure should be dropped from the model. When the Stroop measure is removed, the step 2 model is significant ($AdjR^2$=0.562, p<0.001) and BPVS is the only significant predictor at step 2 (Beta=0.251, p=0.044). This fits into the theory because the model echoes the pattern shown by ecology. Also, looking at the output more closely, both ecology and biodiversity at Time 1 are around p=0.080 so they are having relatively the same predictive power. Likewise, digit recall at Time 1 is at p=0.086. When this analysis for evolution is run again without ecology at Time 1, biodiversity is significant, and when the reverse analysis is run, ecology is significant, digit recall remains close to significance. This suggests that ecology and biodiversity are competing in the evolutionary model, yet biodiversity is somehow more fundamental. The marginal significance of digit recall suggests that verbal working memory is also required, presumably for the effect of language to occur, either as a mechanism for the acquisition of new vocabulary, but also as a potential method for maintaining multiple concepts in the mind when trying to coordinate related ideas. Figure 10.6 below illustrates the final regression models for predicting change in all four biological constructs at Time 2.

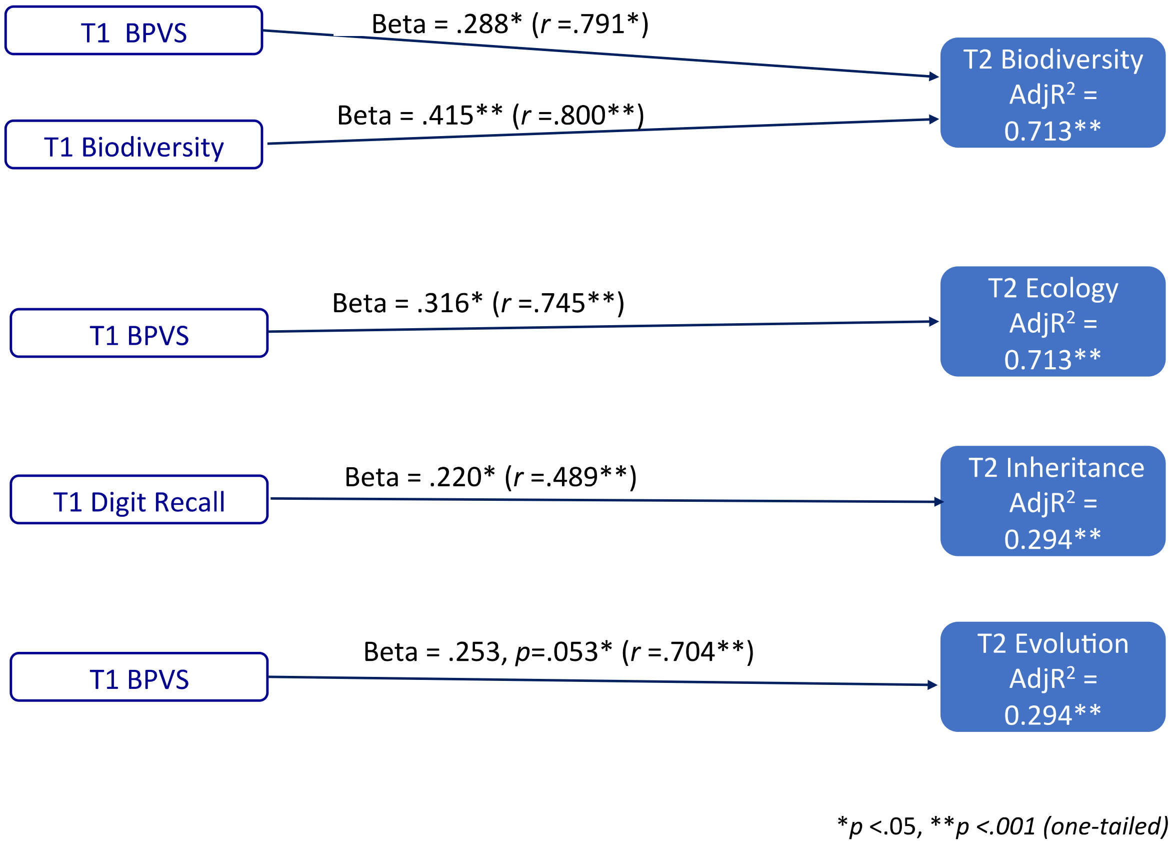


Figure 10.6. Final hierarchical regression models using Time 1 predictors to predict change at Time 2

10.5 Parent Demographics

The influence of parent demographic data on children's conceptual development for each of the biological constructs was examined in more detail. The degree to which parent demographic variables could explain the variance in each of the biological constructs was investigated by running the same two-stage hierarchical regression model as described earlier, by including the parent demographic variables and age in the first stage of the model, and the remaining biological constructs in the second stage. The variables used in the first stage of the model were: age of the child at time of testing, number of adults in the home, number of younger siblings, number of older siblings, preschool attendance, English as a native language, SES, education level of the mother, education level of the father, occupation level of the mother, occupation level of the father, dummy variable for English spoken at home versus other language spoken at home, and dummy variable for English spoken at home versus multiple languages spoken at home (including English). This is a total of 13 variables in the first stage, hence 16 predictors in the second stage. The number of variables relative in these analyses past the point that one is adequately able to use with a sample size of 82, thus these results need to be regarded as preliminary and treated with due caution. Note that this is sample size is lower than the final sample of 129, because not all parents responded to the demographic questionnaire and full data from only a total of 82 parents were available. This sample might affect the results for parameters included in the models presented earlier, therefore, as the sample sizes are not identical.

10.5.1 Preliminary models for Time 1

10.5.1.1 Hierarchical Regressions

In order to examine the influence of certain demographic variables on children's performance on the biological constructs, a two stage hierarchical regression was conducted. The variables included in the first stage of the model are described above. In the second stage of the model, the composite mean scores for biodiversity, ecology, inheritance, and evolution were included. Tables 10.13-10.16 below show the results from the preliminary models for all four biological constructs. All models at step 2 were significant and strong with AdjR^2>0.63 for biodiversity, ecology, and evolution, and AdjR^2>0.43 for inheritance.

Table 10.13. Preliminary demographic only regression model predicting Biodiversity at Time 1

Biodiversity Time 1	Beta	Sig
Model 1: AdjR^2=0.525, p<0.001		
Age	0.712	**<0.001**
No. of adults	0.072	0.394
Older children	0.075	0.390
Younger children	0.136	0.120
Preschool attendance	-0.059	0.463
Free school meals	-0.071	0.406
English native	-0.162	0.077
Education: mother	-0.030	0.795
Education: father	-0.013	0.927
Occupation: mother	0.103	0.371
Occupation: father	-0.066	0.627
English only vs bilingual at home	-0.036	0.699
English vs other language at home	-0.044	0.642
Model 2: AdjR^2=0.766, p<0.001		
Age	0.424	**<0.001**
No. of adults	0.023	0.698
Older children	0.070	0.254
Younger children	0.061	0.329
Preschool attendance	-0.057	0.320

Free school meals	0.034	0.585
English native	-0.057	0.395
Education: mother	-0.200	**0.022**
Education: father	-0.009	0.925
Occupation: mother	0.108	0.188
Occupation: father	0.062	0.528
English only vs bilingual at home	0.015	0.824
English vs other language at home	-0.069	0.307
Ecology Time 1	0.545	**<0.001**
Inheritance Time 1	-0.182	**0.020**
Evolution Time 1	0.196	**0.045**

Table 10.14. Preliminary demographic only regression model predicting ecology at Time 1

Ecology Time 1	Beta	Sig
Model 1: AdjR^2=0.197, p<0.001		
Age	0.449	**<0.001**
No. of adults	0.100	0.363
Older children	-0.017	0.877
Younger children	0.104	0.358
Preschool attendance	0.004	0.970
Free school meals	-0.151	0.178
English native	-0.156	0.188
Education: mother	0.222	0.147
Education: father	-0.088	0.629
Occupation: mother	0.003	0.986
Occupation: father	-0.117	0.508
English only vs bilingual at home	-0.118	0.326
English vs other language at home	-0.017	0.889
Model 2: AdjR^2=0.703, p<0.001		
Age	-0.268	**0.007**

No. of adults	0.038	0.570
Older children	-0.069	0.315
Younger children	-0.050	0.474
Preschool attendance	0.084	0.190
Free school meals	-0.081	0.241
English native	0.046	0.539
Education: mother	0.229	**0.019**
Education: father	-0.054	0.631
Occupation: mother	-0.134	0.148
Occupation: father	-0.071	0.522
English only vs bilingual at home	-0.056	0.443
English vs other language at home	0.040	0.596
Biodiversity Time 1	0.692	**<0.001**
Inheritance Time 1	0.370	**<0.001**
Evolution Time 1	0.148	0.182

Table 10.15. Preliminary demographic regression model predicting inheritance at Time 1

Inheritance Time 1	Beta	Sig
Model 1: AdjR^2=0.123, p<0.001		
Age	0.378	**0.001**
No. of adults	0.031	0.783
Older children	-0.020	0.863
Younger children	0.091	0.437
Preschool attendance	-0.072	0.510
Free school meals	-0.006	0.956
English native	-0.147	0.235
Education: mother	-0.045	0.779
Education: father	-0.115	0.544
Occupation: mother	0.138	0.378
Occupation: father	0.096	0.604
English only vs bilingual at home	-0.094	0.453
English vs other language at home	-0.104	0.415

Model 2: AdjR^2=0.434, p<0.001		
Age	0.231	0.099
No. of adults	-0.007	0.937
Older children	0.011	0.907
Younger children	0.033	0.734
Preschool attendance	-0.078	0.380
Free school meals	0.099	0.298
English native	-0.047	0.647
Education: mother	-0.266	**0.050**
Education: father	-0.089	0.563
Occupation: mother	0.158	0.218
Occupation: father	0.210	0.169
English only vs bilingual at home	-0.023	0.823
English vs other language at home	-0.131	0.208
Biodiversity Time 1	-0.441	**0.020**
Ecology Time 1	0.705	**<0.001**
Evolution Time 1	0.255	0.096

Table 10.16. Preliminary demographic only regression model predicting evolution at Time 1

Evolution Time 1	Beta	Sig
Model 1: AdjR^2=0.460, p<0.001		
Age	0.570	**<0.001**
No. of adults	0.001	0.993
Older children	0.054	0.562
Younger children	0.177	0.058
Preschool attendance	-0.087	0.308
Free school meals	-0.122	0.185
English native	-0.239	**0.015**
Education: mother	0.205	0.103
Education: father	0.119	0.425

	Occupation: mother	0.095	0.440
	Occupation: father	-0.240	0.101
	English only vs bilingual at home	-0.015	0.876
	English vs other language at home	0.078	0.437
Model 2: AdjR^2=0.632, p<0.001			
	Age	0.205	0.069
	No. of adults	-0.045	0.547
	Older children	0.037	0.631
	Younger children	0.101	0.196
	Preschool attendance	-0.058	0.420
	Free school meals	-0.071	0.357
	English native	-0.137	0.099
	Education: mother	0.181	0.101
	Education: father	0.158	0.202
	Occupation: mother	0.040	0.702
	Occupation: father	-0.214	0.081
	English only vs bilingual at home	0.033	0.687
	English vs other language at home	0.112	0.184
	Biodiversity Time 1	0.309	**0.045**
	Ecology Time 1	0.184	0.182
	Inheritance Time 1	0.166	0.096

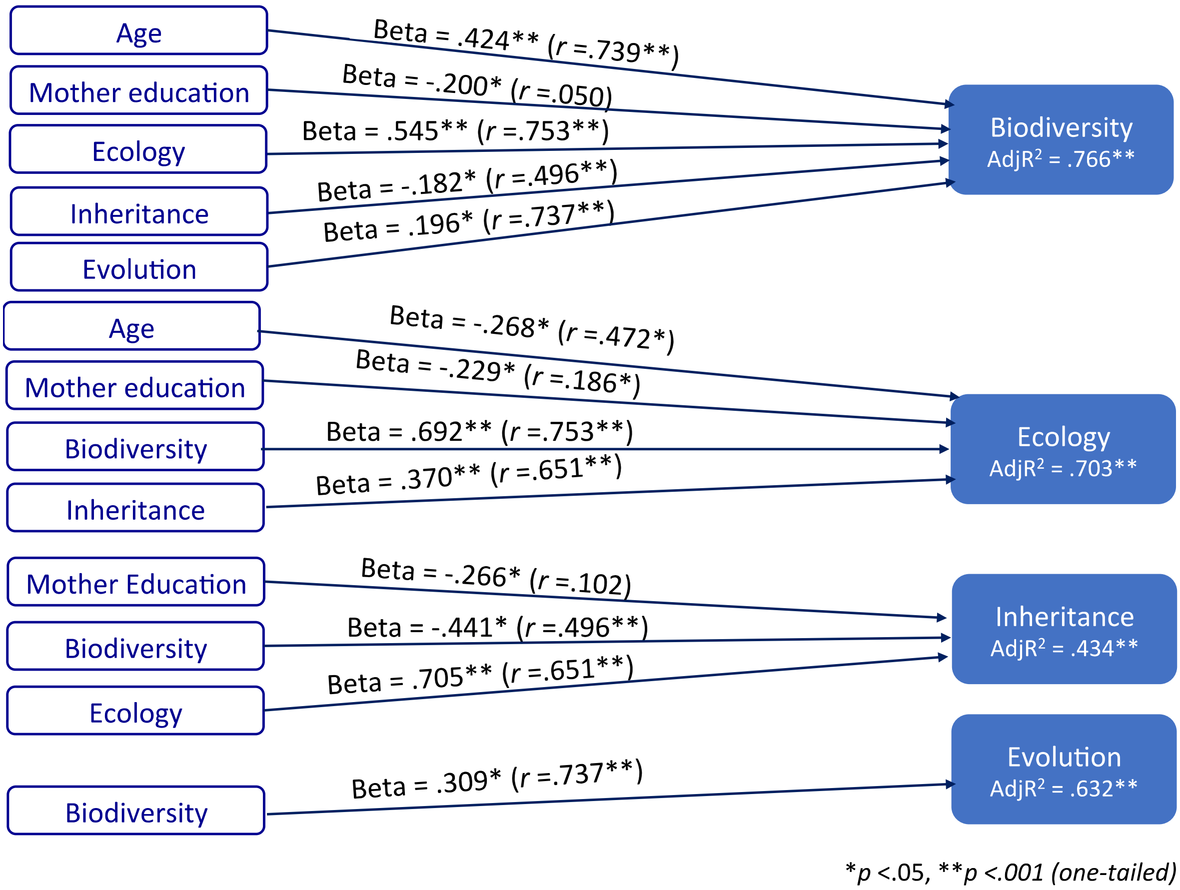


Figure 10.7. Preliminary hierarchical regression models for Time 1 biological constructs

Figure 10.7 suggests that out of all the parent demographic variables, only mothers education level is a significant predictor at Time 1, however that this is negative for biodiversity and inheritance, and positive for ecology. The change in the direction of the significant effect of mother's education level may be down to the unstable effects of having too many predictor variables in the model. This has ultimately led to unstable marginal effects possibly as a function of the smaller pool of participants and the results, therefore, should be treated with caution. With regards to the remaining significant predictors, age is significant for biodiversity and ecology suggesting that experience is significantly contributing to children's knowledge in these two areas. The pattern of findings seem to be different from earlier models (Figures 10.1 and 10.8) but the importance of biodiversity and ecology constructs seems to remain.

10.5.2 Preliminary models for Time 2

10.5.2.1. Hierarchical Regressions

A two-stage hierarchical regression was conducted as above, for Time 2 using the same demographic variables in step 1 included at Time 1, described above. In the second stage of the model, the composite mean scores for biodiversity, ecology, inheritance, and evolution were included. The results of these preliminary analyses all revealed significant regression models at Time 2 ($AdjR^2$>0.69 for biodiversity, ecology, and evolution, and $AdjR^2$>0.49 for inheritance).

The results from Tables 10.17-10.20 below suggest that parent demographic variables seem to have more of an influence at Time 2 than at Time 1. As well as mother's level of education being significant, father's education level, the number of adults and young children in the home, and father's occupation are also significant predictors. However as with the Time 1 models, the direction of the significant relationship varies, for example the number of adults is a significant negative predictor of biodiversity but a significant positive predictor for evolution. Similarly, in the ecology model, father's education is a significant negative predictor whereas father's occupation is a significant positive predictor. This seems unlikely given how related father's education and occupation are and once again, the findings may be a result of the smaller sample size in these analyses relative to the number of parameters in the model. Given this issue, these findings must be interpreted with caution.

The parent demographics seem to significantly contribute to the ecology model the most, with three predictors, whereas the remaining constructs only have one significant parent variable. At Time 2, it also seems as though evolution is a main predictor for the other biological constructs. Note also that a shift in the ecology model at Time 1 with mother's education level to father's education and occupation level at Time 2 suggests that these two variables may essentially be measuring the same thing and that there may be potential for these variables to be collapsed. This will be explored further on.

Interestingly, English as a native language is a significant predictor at step 1 for biodiversity and ecology, and marginally significant for evolution, but drops out at step 2. This would suggest that English as a native language is to some extent a weaker predictor of biological constructs (excluding inheritance) than BPVS, which is superseded by expressive language as shown in Figures 10.3 and 10.4.

With regards to SES, preschool attendance, and both dummy variables assessing language spoken in the home, these variables did not significantly predict any variance in any model (either at step 1 or 2) across Time 1 or Time 2. This consistent non-significant finding would suggest that these variables should be excluded from the model so as to reduce noise in the data, as was done in earlier models (see Figures 10.1 and 10.2). This left a total of six predictors for step 1, and 10 predictors in total for step 2. With a sample of 82, the model is still somewhat underpowered, but stronger than the weaker models presented in the preliminary analyses here. The extraneous variables excluded were: the number of older children living at home, preschool education, SES, both the dummy variables for language(s) spoken in the home, and mother's occupation. With these variables now identified, the regression analyses for Time 1 and for Time 2 were repeated in the same way as described above.

Table 10.17. Preliminary demographic only regression model predicting biodiversity at Time 2

Biodiversity Time 2	Beta	Sig
Model 1: AdjR^2=0.594, p<0.001		
Age	0.658	**<0.001**
No. of adults	-0.111	0.164
Older children	-0.042	0.612
Younger children	-0.009	0.913
Preschool attendance	0.031	0.683
Free school meals	0.043	0.603
English native	-0.198	**0.028**
Education: mother	0.040	0.712
Education: father	-0.108	0.403
Occupation: mother	0.020	0.856
Occupation: father	0.179	0.160
English only vs bilingual at home	-0.029	0.739
English vs other language at home	-0.102	0.266
Model 2: AdjR^2=0.755, p<0.001		
Age	0.201	**0.038**
No. of adults	-0.185	**0.004**
Older children	-0.035	0.589
Younger children	0.060	0.373
Preschool attendance	-0.009	0.887
Free school meals	0.059	0.368

English native	-0.047	0.521
Education: mother	-0.051	0.588
Education: father	0.130	0.233
Occupation: mother	0.054	0.539
Occupation: father	-0.035	0.738
English only vs bilingual at home	-0.002	0.978
English vs other language at home	-0.010	0.893
Ecology Time 2	0.305	**0.008**
Inheritance Time 2	0.076	0.400
Evolution Time 2	0.335	**0.003**

Table 10.18. Preliminary demographic only regression model predicting ecology at Time 2

Ecology Time 2	Beta	Sig
Model 1: AdjR^2=0.584, *p*<0.001		
Age	0.692	**<0.001**
No. of adults	0.091	0.260
Older children	-0.047	0.574
Younger children	-0.163	0.059
Preschool attendance	0.082	0.288
Free school meals	0.032	0.704
English native	-0.256	**0.006**
Education: mother	0.084	0.446
Education: father	-0.354	**0.009**
Occupation: mother	-0.076	0.503
Occupation: father	0.364	**0.006**
English only vs bilingual at home	-0.011	0.903
English vs other language at home	-0.100	0.282
Model 2: AdjR^2=0.708, *p*<0.001		
Age	0.301	**0.004**
No. of adults	0.097	0.184
Older children	-0.032	0.656

Younger children	-0.151	**0.039**
Preschool attendance	0.070	0.284
Free school meals	0.024	0.743
English native	-0.139	0.080
Education: mother	0.063	0.536
Education: father	-0.245	**0.037**
Occupation: mother	-0.082	0.392
Occupation: father	0.238	**0.034**
English only vs bilingual at home	0.020	0.788
English vs other language at home	-0.027	0.731
Biodiversity Time 2	0.364	**0.008**
Inheritance Time 2	-0.037	0.708
Evolution Time 2	0.266	**0.031**

Table 10.19. Preliminary demographic only regression model predicting inheritance at Time 2

Inheritance Time 2	Beta	Sig
Model 1: AdjR^2=0.380, p<0.001		
Age	0.461	**<0.001**
No. of adults	0.020	0.835
Older children	0.064	0.533
Younger children	-0.067	0.516
Preschool attendance	0.115	0.224
Free school meals	-0.134	0.196
English native	-0.130	0.235
Education: mother	0.465	**0.001**
Education: father	-0.349	**0.032**
Occupation: mother	-0.066	0.636
Occupation: father	0.221	0.161
English only vs bilingual at home	0.013	0.905
English vs other language at home	-0.136	0.229
Model 2: AdjR^2=0.496, p<0.001		

Age	0.125	0.376
No. of adults	-0.015	0.878
Older children	0.064	0.491
Younger children	-0.057	0.555
Preschool attendance	0.107	0.212
Free school meals	-0.118	0.211
English native	-0.033	0.754
Education: mother	0.425	**0.001**
Education: father	-0.219	0.159
Occupation: mother	-0.067	0.596
Occupation: father	0.103	0.495
English only vs bilingual at home	0.050	0.614
English vs other language at home	-0.060	0.565
Biodiversity Time 2	0.156	0.400
Ecology Time 2	-0.064	0.708
Evolution Time 2	0.439	**0.006**

Table 10.20. Preliminary demographic only regression models predicting evolution at Time 2

Evolution Time 2	Beta	Sig
Model 1: AdjR^2=0.443, p<0.001		
Age	0.631	**<0.001**
No. of adults	0.133	0.155
Older children	0.008	0.937
Younger children	-0.043	0.659
Preschool attendance	0.018	0.843
Free school meals	-0.047	0.632
English native	-0.188	0.073
Education: mother	0.089	0.486
Education: father	-0.309	**0.045**
Occupation: mother	-0.016	0.903
Occupation: father	0.258	0.085
English only vs bilingual at home	-0.075	0.467

English vs other language at home	-0.153	0.155
Model 2: AdjR^2=0.690, p<0.001		
Age	0.031	0.782
No. of adults	0.149	**0.045**
Older children	0.022	0.767
Younger children	0.025	0.745
Preschool attendance	-0.050	0.462
Free school meals	-0.038	0.610
English native	0.004	0.964
Education: mother	-0.078	0.460
Education: father	-0.068	0.581
Occupation: mother	0.015	0.882
Occupation: father	0.019	0.871
English only vs bilingual at home	-0.063	0.415
English vs other language at home	-0.045	0.583
Biodiversity Time 2	0.425	**0.003**
Ecology Time 2	0.283	**0.031**
Inheritance Time 2	0.271	**0.006**

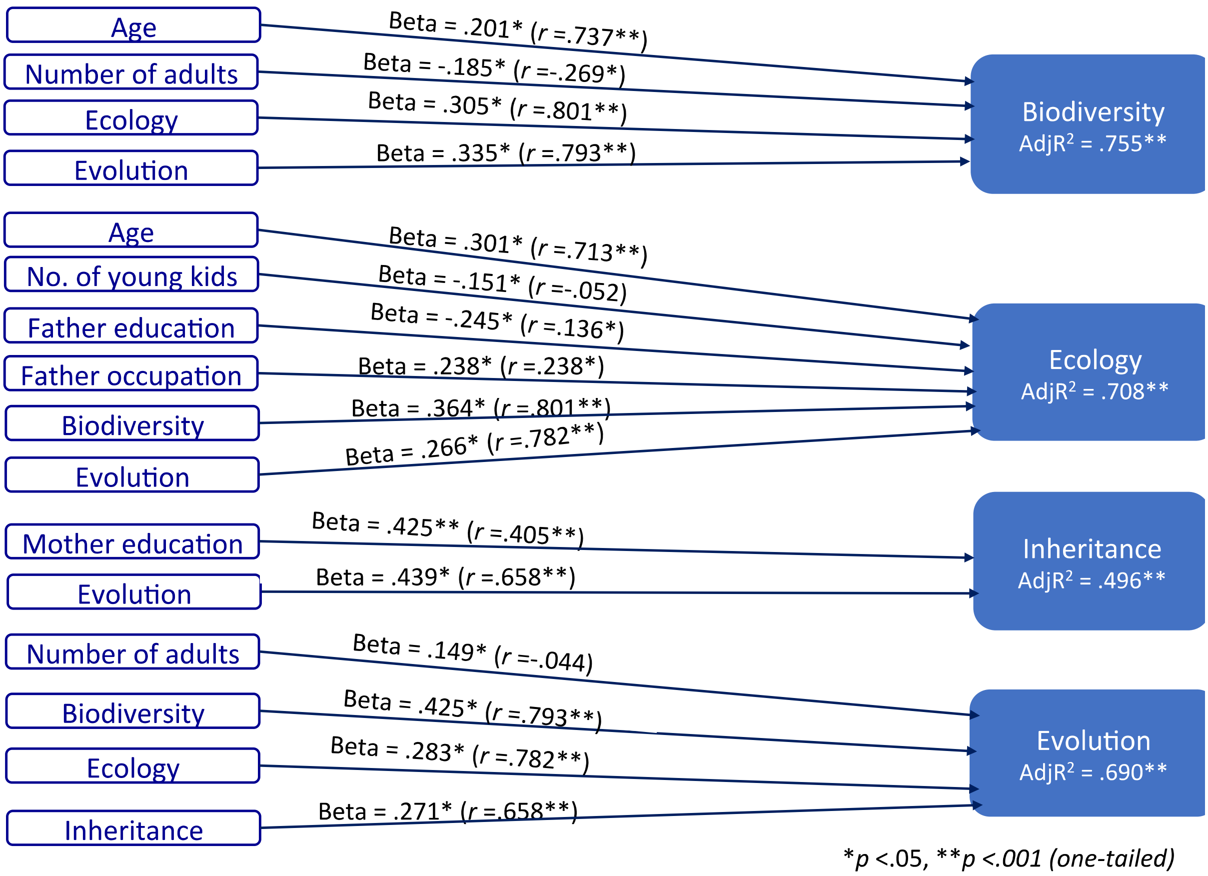


Figure 10.8. Preliminary hierarchical regression models for Time 2 biological constructs

10.5.2.1.1 Final models of significant demographic variables predicting biological constructs Time 1

After removing the extraneous demographic parameters from the models presented in Figures 10.8 and 10.9, a total of 10 predictors were used in the final models: Age, number of younger siblings, English native, mother's education level, father's education level, father's occupation level, number of adults in the home, and three of the remaining biological constructs to the one being modelled in the analysis. The hierarchical regressions produced significant and strong models ($AdjR^2$>0.69 for biodiversity, ecology, and evolution, and $AdjR^2$>0.43 for inheritance).

Biodiversity: AdjR^2=0.539 for step 1; AdjR^2=0.763 for step 2 (p<0.001). Ecology: AdjR^2=0.225 for step 1; AdjR^2=0.692 for step 2 (p<0.001). Inheritance: AdjR^2=0.174 (p<0.050) for step 1; AdjR^2=0.430 for step 2 (p<0.001). Evolution: AdjR^2=0.468 for step 1; AdjR^2=0.643 for step 2 (p<0.001).

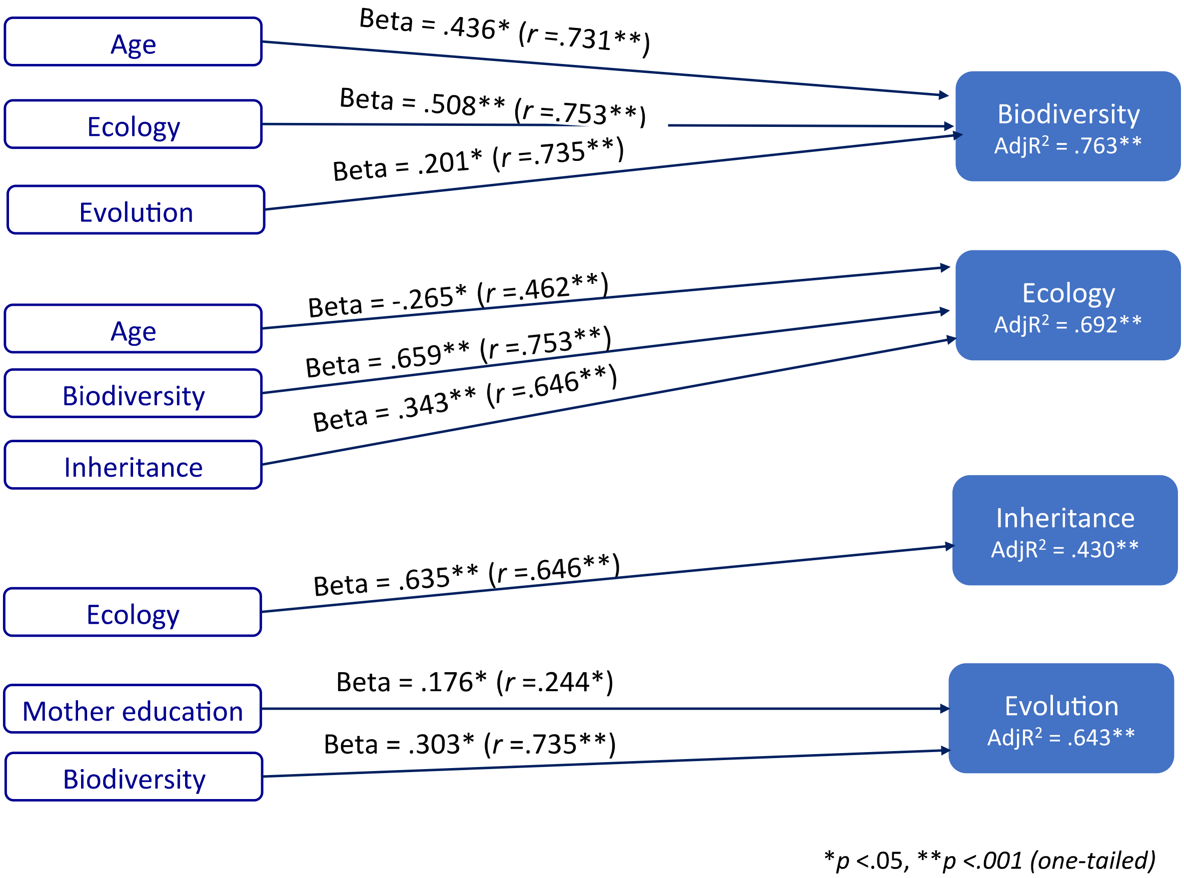


Figure 10.9. Final regression models with extraneous variables removed at Time 1

When the extraneous variables are removed in the Time 1 models, the models largely remain the same. The key difference is the fact that mother's education is no longer a significant predictor for biodiversity, ecology, and inheritance, and instead it *is* a significant predictor for the evolution model, where previously in the preliminary models, it was not. Removing noise from the data suggests that in actuality, parent demographics have very little influence over children's biological knowledge, aside from mother's education level which seems to be an important predictor for children's learning and development. However, this does not explain why the locus of influence of the biodiversity and age parameters shifts around from having a significant negative influence to a significant positive influence or vice versa. It may be that the influences are marginal and shift around according to what other parameters are included in the model, which would be consistent with the preliminary demographic models, and also the influence of BPVS on the Time 2 constructs described in section 10.4.1. This is most likely a function of too many variables relative to the small sample size and these findings while interpreted with caution, may also not be comparable with the cognitive models described in sections 10.2 and 10.3.

10.5.2.1.2 Final models of significant demographic variables predicting biological constructs Time 2

The same analyses were conducted for Time 2 demographic only models using the same consistently significant demographic variables from the preliminary analyses at Time 1 and Time 2. All models were significant at $p>0.001$ and stronger than they were at Time 1, with $AdjR^2$ 0.704 for biodiversity, ecology, and evolution. The final inheritance model was $AdjR^2=0.704$ for step 2 ($p<0.001$). A summary of the final regression models is displayed in Figure 10.10 below.

Biodiversity: $AdjR^2=0.590$ for step 1; $AdjR^2=0.774$ for step 2 ($p<0.001$). Ecology: $AdjR^2=0.561$ for step 1; $AdjR^2=0.714$ for step 2 ($p<0.001$). Inheritance: $AdjR^2=0.379$ ($p<0.050$) for step 1; $AdjR^2=0.493$ for step 2 ($p<0.001$). Evolution: $AdjR^2=0.459$ for step 1; $AdjR^2=0.704$ for step 2 ($p<0.001$).

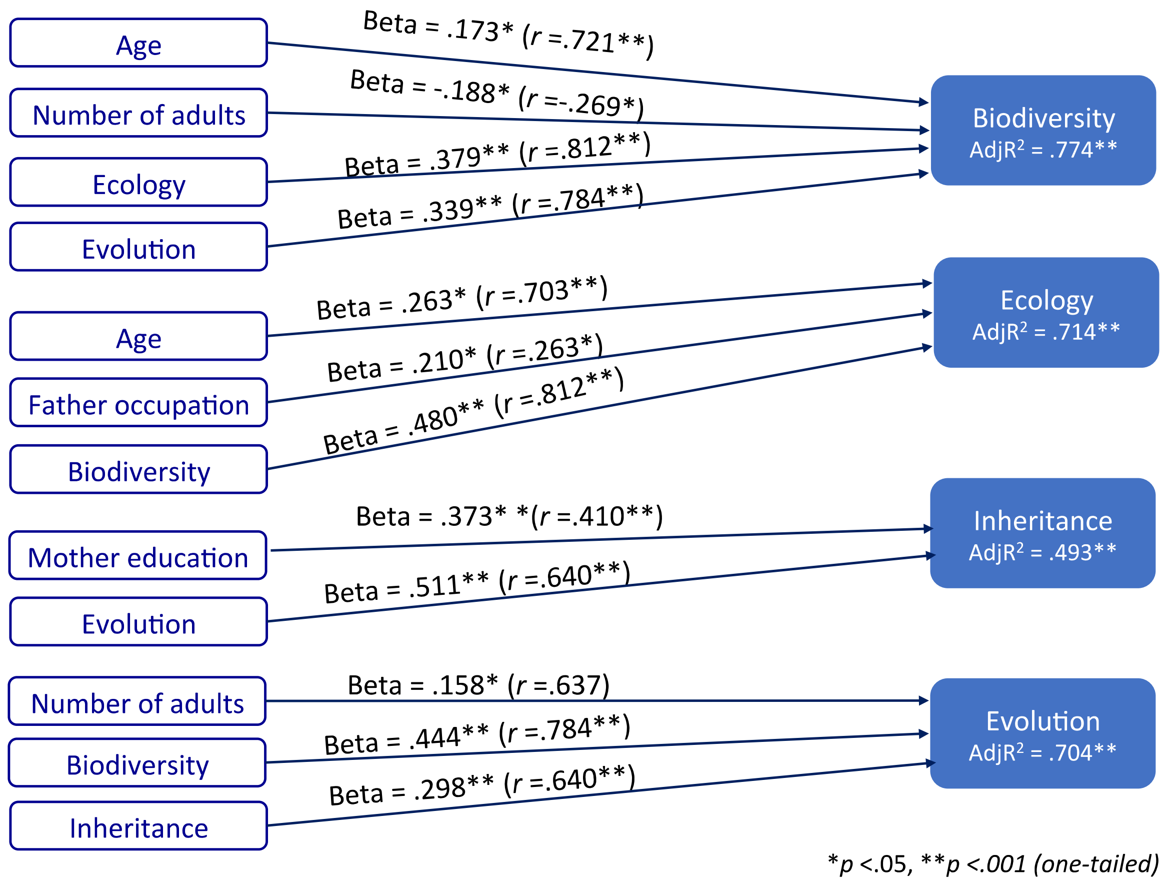


Figure 10.10. Final regression models with extraneous variables removed at Time 2

At Time 2 it seems that the models vary from Time 1 because the status of evolution changes, where is now a significant positive predictor for inheritance, replacing ecology at Time 1. This follows a similar pattern to the regression models presented in sections 10.2 and 10.3 of the cognitive parameters, where it was observed that ecology and evolution might have a reciprocal relationship with regards to influencing children's inheritance knowledge. The biodiversity model shows no change in significant predictors, but for ecology however, there are far fewer significant parent variables with the only remaining variable being father's occupation level. These final models suggest that again, parent demographic variables have little influence over children's biological knowledge, but the key predictors that do have a significant contribution are: mother's education level, father's occupation level, and the number of adults in the home.

However the problem of too many variables relative to the sample size remains. One way to account for this was to create a composite variable of mother and father levels of education because they might be measuring the same thing. As would mother and father levels of occupation. This will be explored further on. A puzzling finding was the significant predictive influence of the number of adults in the home for biodiversity and evolution models at Time 2. This also needed to be explored in more detail.

10.5.3 Number of adults

The number of adults in the home was a significant variable in a couple of the models (see Figure 10.10), although the influence of this parameter was sometimes negative and sometimes positive. In order to examine the effect of the number of adults in the home on children's biological knowledge in more detail, plots were constructed to reveal the pattern of this influence at Time 1 and at Time 2. These are shown in Figures 10.11-10.14 below:

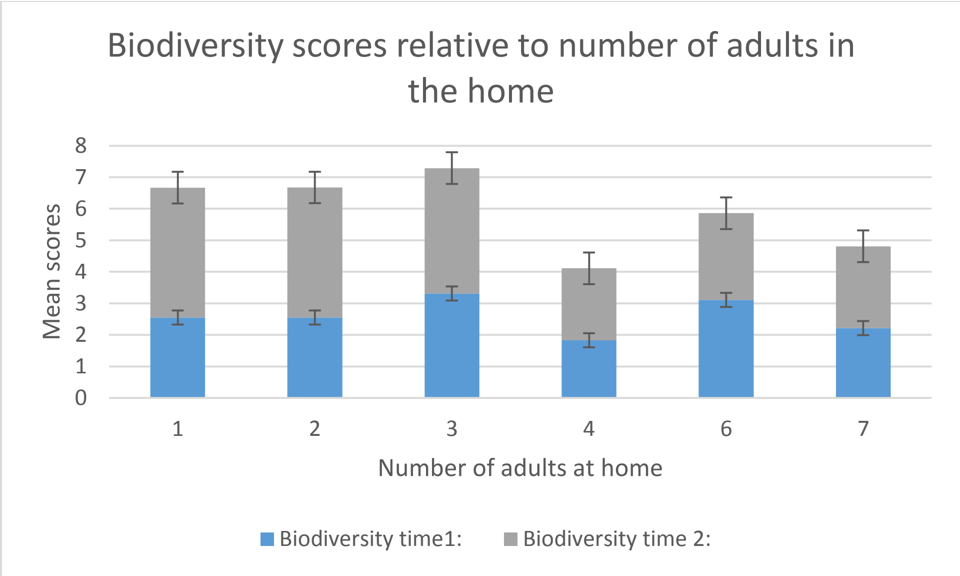


Figure 10.11. Relationship the number of adults at home with biodiversity at Time 1 and two

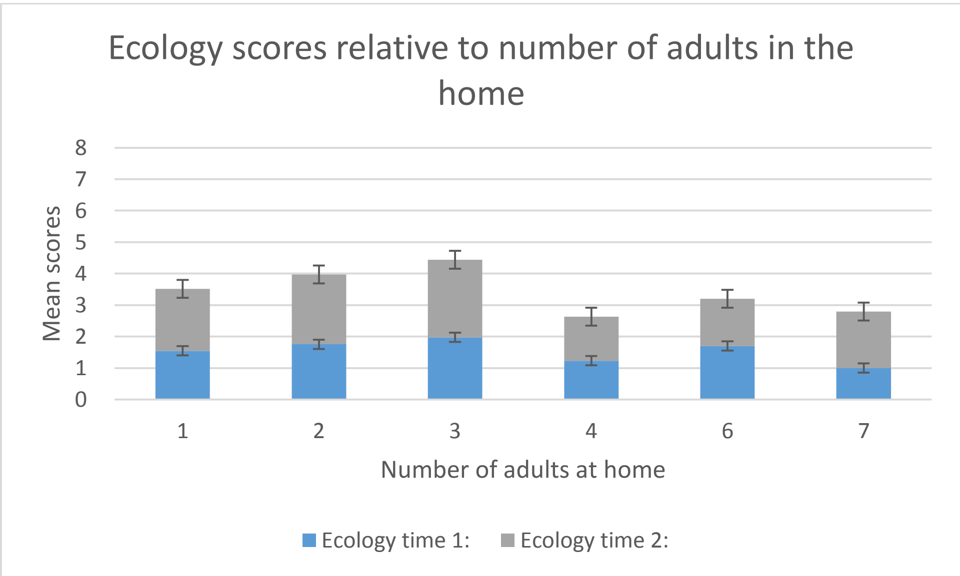


Figure 10.12. Relationship with number of adults at home with ecology at Time 1 and two

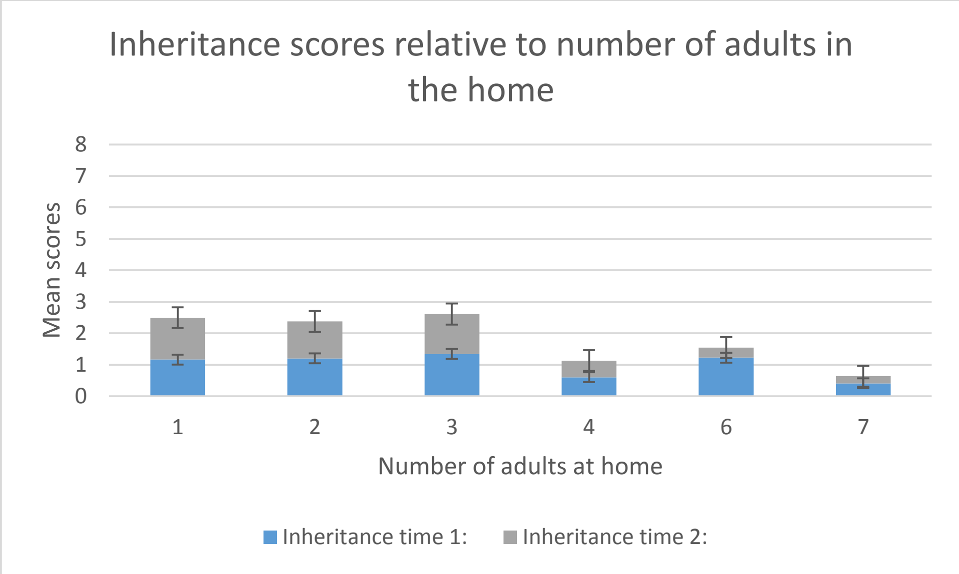


Figure 10.13. Relationship with number of adults at home and inheritance at Time 1 and two

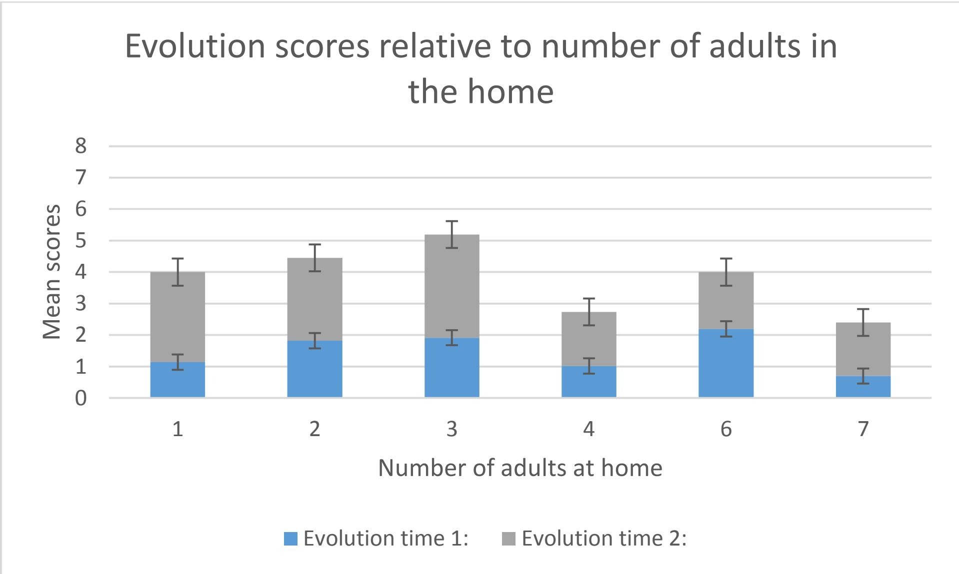


Figure 10.14. Relationship with number of adults in the home and evolution at Time 1 and Time 2

These graphs generally suggest that three adults in the home is the optimal number, however this is not the case for inheritance. The third adult was typically reported as a grandmother. It may be that the presence of an extra adult may provide the child with more resources including time, knowledge, outdoor activities, museum trips etc., which might contribute toward the increase in children's biological knowledge. Interestingly, more than three adults in the home does not provide an extra advantage, rather children's performance seems to be similar to that of having one or two adults in the home. Reasons for why more than four adults in the home might not have an increasingly optimal effect on children's biological knowledge are discussed in Chapter 11.

A thing to consider with these plots is that only two children out of the whole sample reported either 6 or 7 adults as being in the home. For these reasons it was thought best to include these children in the group with 4 adults, and simply collapse the group to "4+ adults in the home" and run any future analysis using this new variable.

10.6 Final composite hierarchical regression models

As discussed above (see Figure 10.10) variables on parent levels of education and occupation seemed to have the most influence over children's biological knowledge. Due to the number of multiple variables included in the model however, an attempt was made to create a composite variable of mother and father levels of education and occupation.

10.6.1 Parent occupation and education

The relationship between father and mother levels of occupation and education on biological knowledge were harder to interpret. There seems to be some shift between mother and father level of education as seen in the preliminary parent demographic only models, which suggests that these two variables are correlated. Correlations of all parent education and occupation levels are shown below in Table 10.21. As can be seen, all four variables are significantly correlated with each other:

Table 10.21. Correlations between parent education and occupation levels

		Education level of mother of child	Education level of father of child	Occupation level of mother	Occupation level of father
Education level of mother of child	Pearson Correlation	1	0.595**	0.616**	0.488**
	N	104	90	104	90
Education level of father of child	Pearson Correlation		1	0.381**	0.788**
	N		93	91	88
Occupation level of mother	Pearson Correlation			1	0.415**
	N			109	92
Occupation level of father	Pearson Correlation				1
	N				94

*p<0.05; **p<0.001

To account for this high degree of correlation, Cronbach's alpha was calculated for all of these four items. The alpha value came out at 0.786. This is a high degree of reliability suggesting these four items are related. However, when viewing the item-total statistics, dropping mother's level of occupation as a variable increases in the alpha value to 0.820 as shown in Table 10.22.

Table 10.22. Item-total statistics of Cronbach's alphas for parent education and occupation levels

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Education level of mother of child	6.950	20.379	0.711	0.531	0.726
Education level of father of child	6.680	18.886	0.677	0.658	0.712
Occupation level of mother	6.420	14.199	0.522	0.406	0.820
Occupation level of father	5.140	15.075	0.674	0.643	0.690

This suggests that although mothers are often equivalently educated as the fathers, many of the mothers do not have an equal level of occupation to the fathers. This may be because in the current sample, many mothers have reported their choice to become homemakers after having young children or taking short-term career breaks. For these reasons mother's occupation was excluded from a composite measure, which ultimately consisted of an average of mother's level of education, father's level of education, and father's level of occupation that was created for each child. The new variable "parent edu/occ" was included in the same demographic only regression models, before also running the composite regression models (general cognitive ability and parent demographic models) as above to see what affect this would have and whether interpretation of these models would be easier.

These new models show no change in the demographic only regression models (excluding extraneous variables) with the new 'parent edu/occ' variable included at Time 1, but contrastingly the model changes for ecology at Time 2 where 'father's occupation' drops out as a significant variable. Also for evolution at Time 1, 'mother education' is replaced with 'parent edu/occ' composite variable. At Time 2, the models for ecology and inheritance change in a similar way.

10.6.1.1 Time 1 regression models with new parent composite variable

Biodiversity: AdjR^2=0.542 for step 1; AdjR^2=0.717 for step 2 (p<0.001). Ecology: AdjR^2=0.244 for step 1; AdjR^2=0.651 for step 2 (p<0.001). Inheritance: AdjR^2=0.200 for step 1; AdjR^2=0.396 for step 2 (p<0.001). Evolution: AdjR^2=0.464 for step 1; AdjR^2=0.614 for step 2 (p<0.001). The results of the models can be viewed in the appendices (A.6). The final models for each biological construct have been summated in Figure 10.15 below.

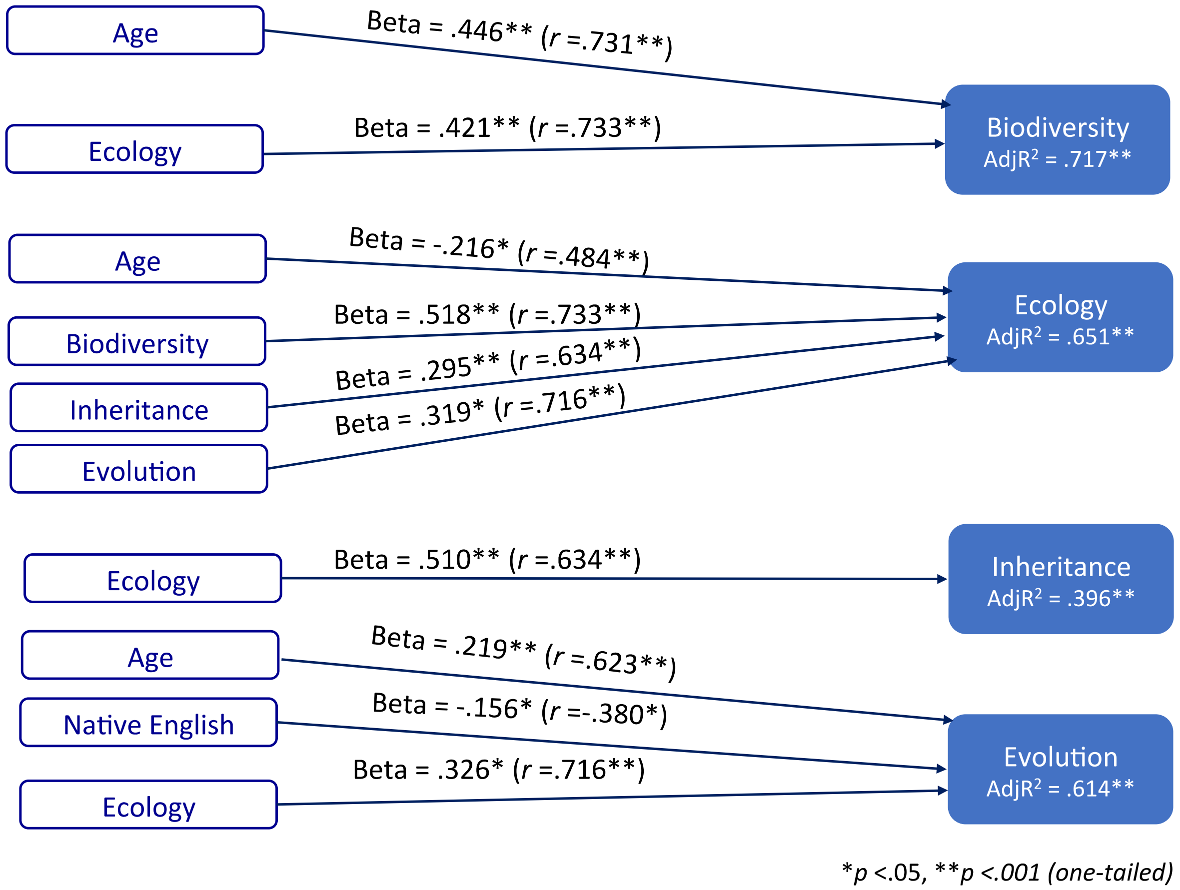


Figure 10.15. Exploratory regression models using only parent demographic data as predictors for biological constructs at (excluding extraneous variables) at Time 1 with new parent education/occupation composite variable.

This model suggests that age is significant predictor for all biological constructs aside from inheritance, indicating that experience is a crucial factor in the development of biological knowledge. Ecological knowledge also seems to be core in terms of driving change in related biological areas. Interestingly the only demographic variable that significantly predicts a biological construct is being a native English speaker for evolutionary knowledge. This is a curious finding and may relate to children being able to sufficiently articulate their knowledge during the biological task, which being a native speaker of English, would be easier to do.

10.6.1.2 Time 2 regression models with new parent composite variable

Biodiversity: AdjR^2=0.577 for step 1; AdjR^2=0.780 for step 2 (p<0.001). Ecology: AdjR^2=0.508 for step 1; AdjR^2=0.713 for step 2 (p<0.001). Inheritance: AdjR^2=0.301 for step 1; AdjR^2=0.441 for step 2 (p <0.001). Evolution: AdjR^2=0.409 for step 1; AdjR^2=0.701 for step 2 (p<0.001). The results of the models can be viewed in the appendices (A.7). The significant predictors for the final models for each biological construct have been summarised in Figure 10.16.

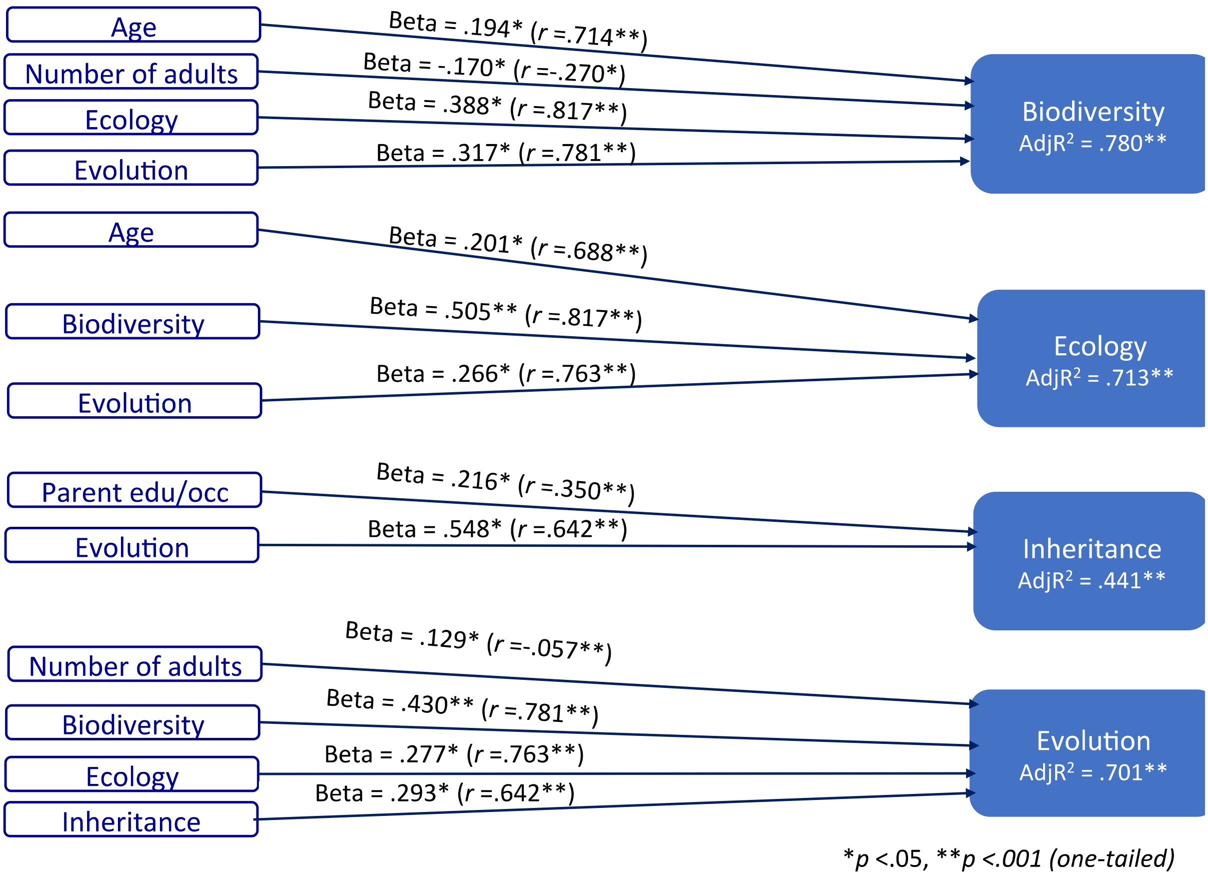


Figure 10.16. Exploratory regression models using only parent demographic data as predictors for biological constructs (excluding extraneous variables) at Time 2 with new parent education/occupation composite variable.

The models at Time 2 suggest that the only two significant parent predictors of children's knowledge at Time 2 are the number of adults in the home and the new parent edu/occ composite variable. This new variable only significantly predicts children's inheritance knowledge at Time 2, which may be because parents who are better educated might be able to communicate the complexity of reproduction more effectively to children. On the whole, this model generally remains the same as the demographic only regression models described earlier.

10.6.2 Composite models Time 1

Having removed the extraneous variables from the preliminary parent demographic models and creating a composite parent education/occupation variable, the significant variables were now included in the final regression models, alongside the significant general cognitive ability measures to predict each biological construct at Time 1 and Time 2, reported below. Note again that this sample size is 82 as not all parents responded to the demographic questionnaire, which might affect the results for the parameters included in the models presented earlier as the sample sizes are not identical.

The final composite regression models for Time 1 suggest that ecology is central to understanding other related areas to biology. Interestingly, when general cognitive abilities are included in the model, the parent variables are no longer significantly contributing to any model. This would suggest that cognitive models are in a sense controlling for demographic variation as well as age. Furthermore, the only general cognitive ability that has a significant influence is backward digit recall, which significantly predicts biodiversity at

Time 1. This implies that executive control is required in order to develop concepts around biodiversity. If the full results are viewed (Tables 10.23-10.26) for these models, it can be seen that at step 1 for each of the models, receptive language as measured by the BPVS is a significant predictor, but drops out of the model at stage 2. This, like earlier models, implies that biologically specific language is vital toward driving forward conceptual change. It may be that language acquired through earlier concepts of biodiversity and ecology need to be coordinated, as suggested by the significance of executive control, which would help coordinate fragmented ideas and encourage 'systems thinking', particularly around ecological concepts which lend themselves to this style of scientific thinking. This then may contribute towards change and development of inheritance and evolutionary concepts at Time 1.

If the previous demographic only models are recalled however, English as a native language had much the same type of relationship as BPVS i.e. always a significant predictor at step 1, and dropping out at step 2. In these composite models, English as a native language never reaches the point of significance at step 1, suggesting speaking English as a native language is superseded by children's receptive language ability. If the earlier cognitive models are also recalled for Time 2, expressive language displays a different type of relationship in that, unlike BPVS or native English variables, it remains in the step 2 model for biodiversity. All this would suggest a hierarchical relationship between these three variables in that English as a native language is superseded by children's receptive language ability, which is superseded by children's expressive language ability.

Table 10.23. Composite hierarchical regression model for biodiversity at Time 1

Biodiversity Time 1	Beta	Sig.
Model 1: AdjR^2=0.631 (p<0.001)		
Age	0.130	0.367
BPVS	0.442	**0.003**
Digit recall	0.099	0.288
b/w digit recall	0.286	**0.007**
Number of adults (comp)	0.069	0.309
Number of younger children	0.166	**0.018**
English native	0.011	0.886
Parent edu/occ	-0.124	0.095
Model 2: AdjR^2=0.765 (p<0.001).		
Age	0.125	0.280
BPVS	0.146	0.251
Digit recall	0.072	0.347
b/w digit recall	0.303	**0.001**
Number of adults (comp)	0.040	0.464
Number of younger children	0.090	0.116
English native	0.054	0.403
Parent edu/occ	-0.096	0.106
Ecology time 1	0.417	**<0.001**
Inheritance time 1	-0.029	0.689
Evolution time 1	0.101	0.274

Table 10.24. Composite hierarchical regression model for ecology at Time 1

Ecology Time 1	Beta	Sig.
Model 1: AdjR^2=0.319 (p<0.001)		
Age	-0.029	0.881
BPVS	0.672	**0.001**
Digit recall	0.026	0.839

b/w digit recall	-0.067	0.633
Number of adults (comp)	0.063	0.497
Number of younger children	0.151	0.109
English native	-0.066	0.542
Parent edu/occ	-0.070	0.485
Model 2: AdjR^2=0.669 (p<0.001)		
Age	-0.179	0.194
BPVS	0.161	0.285
Digit recall	-0.115	0.204
b/w digit recall	-0.194	0.070
Number of adults (comp)	0.025	0.702
Number of younger children	-0.028	0.677
English native	0.004	0.958
Parent edu/occ	0.009	0.895
Biodiversity Time 1	0.587	**<0.001**
Inheritance Time 1	0.237	**0.004**
Evolution Time 1	0.336	**0.002**

Table 10.25. Composite hierarchical regression model for inheritance at Time 1

Inheritance Time 1		
	Beta	Sig.
Model 1: AdjR^2=0.273 (p<0.001)		
Age	0.058	0.774
BPVS	0.596	**0.005**
Digit recall	0.080	0.541
b/w digit recall	-0.232	0.113
Number of adults (comp)	-0.040	0.672
Number of younger children	0.112	0.248
English native	-0.075	0.499
Parent edu/occ	-0.036	0.725

Model 2: AdjR^2=0.404 (p<0.001)		
Age	0.054	0.773
BPVS	0.294	0.144
Digit recall	0.049	0.691
b/w digit recall	-0.188	0.192
Number of adults (comp)	-0.065	0.456
Number of younger children	0.036	0.696
English native	-0.021	0.841
Parent edu/occ	-0.016	0.863
Biodiversity Time 1	-0.073	0.689
Ecology Time 1	0.425	**0.004**
Evolution Time 1	0.148	0.318

Table 10.26. Composite hierarchical regression model for evolution at Time 1

Evolution Time 1		
	Beta	Sig.
Model 1: AdjR^2=0.499 (p<0.001)		
Age	0.178	0.289
BPVS	0.327	0.057
Digit recall	0.188	0.086
b/w digit recall	0.039	0.744
Number of adults (comp)	0.020	0.802
Number of younger children	0.165	**0.043**
English native	-0.175	0.061
Parent edu/occ	0.006	0.948
Model 2: AdjR^2=0.646 (p<0.001)		
Age	0.163	0.252
BPVS	-0.034	0.826
Digit recall	0.157	0.093
b/w digit recall	0.040	0.720
Number of adults (comp)	-0.010	0.884

Number of younger children	0.075	0.286
English native	-0.146	0.065
Parent edu/occ	0.053	0.472
Biodiversity Time 1	0.153	0.274
Ecology Time 1	0.360	**0.002**
Inheritance Time 1	0.088	0.318

The results from the final models of the composite regression analyses at Time 1 are summarised in Figure 10.17 shown below:

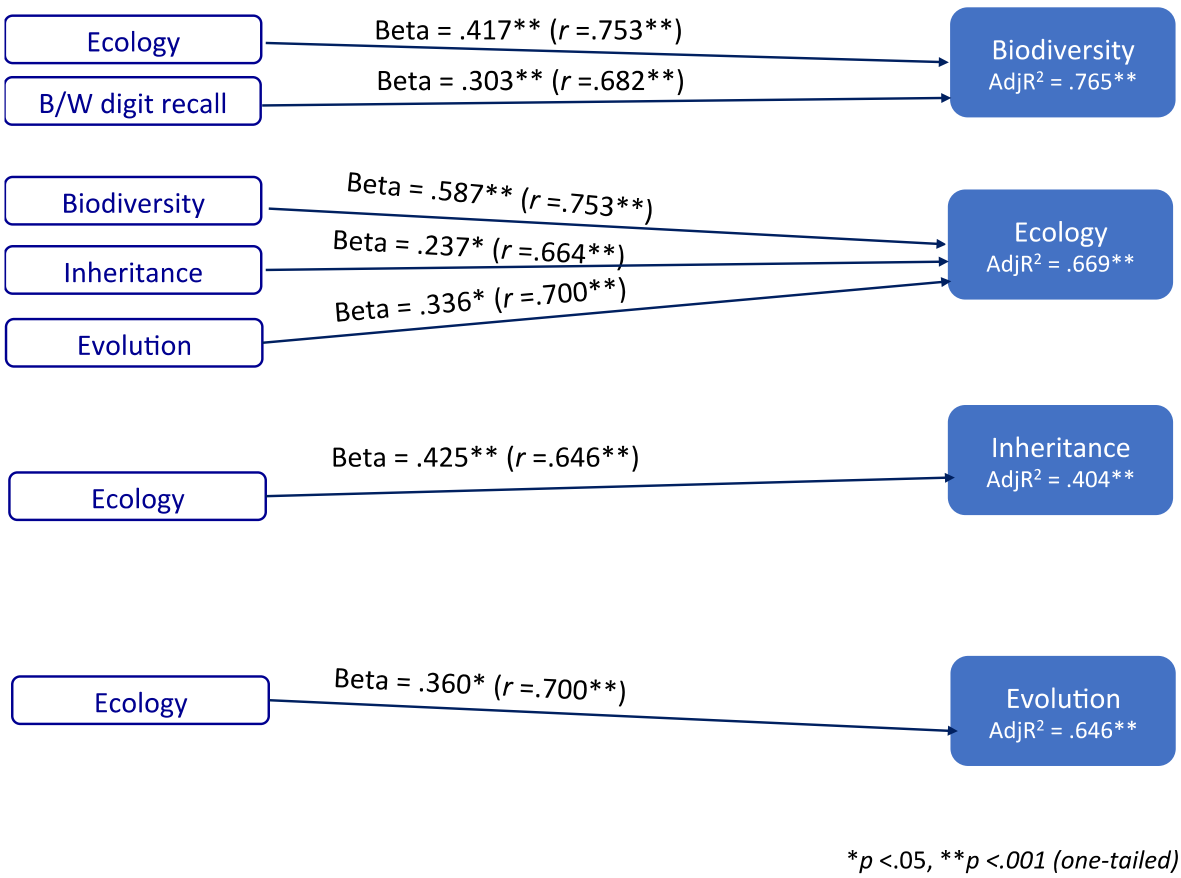


Figure 10.17. Composite hierarchical regression models at Time 1 using significant demographic and cognitive predictors

10.6.3 Composite models Time 2

The models at Time 2 suggest that more parent variables are influential at Time 2 than at Time 1, namely for biodiversity and evolution where the number of adults seems to be key, and the level of parent education/occupation for evolution in particular. The number of adults in the home may be important in terms of the resources children have available to them in the form of time, trips away, books read to them etc. Likewise the same may be true of the number of younger children where the relationship is a negative one perhaps due to a lack of resources. Parent education and occupation levels are also likely to be influential with regards to the knowledge parents are able to pass on to their children and possibly also the number of educational activities, which are all likely to influence children's knowledge.

Additionally, expressive language is the only general cognitive ability measure that significantly predicts knowledge of biodiversity at Time 2. This is interesting because it would confirm the notion described above that expressive language supersedes receptive language, which supersedes English as a native language. This also implies that biologically specific language is key for biodiversity at Time 2, but also highlights the overall importance of language ability in the developmental of biological knowledge in general.

The addition of evolution and ecology as significant predictors also suggests that ideas around biodiversity need to be coordinated and the nature of ecology and evolution aids this coordination, which in turn feeds into the other concepts. The only unique model would be for inheritance (Table 10.29) where the only significant predictor is evolution. At Time 1, it was ecology. This switch may simply refer to the fact that by Time 2 children in all cohorts have better knowledge on evolution than ecology. The results from all the models at Time 2 are presented in Tables 10.27-10.30, and the significant predictors for each biological construct are highlighted in Figure 10.18 below.

Table 10.27. Composite hierarchical regression model for biodiversity at Time 2

Biodiversity Time 2		
	Beta	Sig.
Model 1: AdjR^2=0.726 (p<0.001)		
Age	0.038	0.758
BPVS	0.042	0.779
Digit recall	-0.034	0.682
b/w digit recall	0.055	0.554
Expressive language	0.752	**<0.001**
Number of adults (comp)	-0.089	0.140
Number of younger children	0.075	0.227
English native	-0.034	0.635
Parent edu/occ	-0.026	0.700
Model 2: AdjR^2=0.810 (p<0.001)		
Age	-0.006	0.957
BPVS	-0.059	0.640
Digit recall	-0.059	0.397
b/w digit recall	0.045	0.562
Expressive language	0.481	**0.001**
Number of adults (comp)	-0.134	**0.010**
Number of younger children	0.093	0.078
English native	-0.001	0.991
Parent edu/occ	0.004	0.944
Ecology Time 2	0.347	**<0.001**

Inheritance Time 2	0.005	0.940
Evolution Time 2	0.190	**0.044**

Table 10.28. Composite hierarchical regression model for ecology at Time 2

Ecology Time 2		
	Beta	Sig.
Model 1: AdjR^2=0.572 (p<0.001)		
Age	0.159	0.308
BPVS	0.256	0.175
Digit recall	0.010	0.920
b/w digit recall	0.032	0.783
Expressive language	0.375	0.053
Number of adults (comp)	0.061	0.414
Number of younger children	-0.045	0.558
English native	-0.055	0.535
Parent edu/occ	-0.019	0.823
Model 2: AdjR^2=0.706 (p<0.001)		
Age	0.156	0.229
BPVS	0.215	0.170
Digit recall	0.008	0.923
b/w digit recall	-0.003	0.974
Expressive language	-0.201	0.285
Number of adults (comp)	0.076	0.250
Number of younger children	-0.090	0.174
English native	-0.023	0.753
Parent edu/occ	0.036	0.616
Biodiversity Time 2	0.536	**<0.001**
Inheritance Time 2	-0.055	0.517
Evolution Time 2	0.274	**0.019**

Table 10.29. Composite hierarchical regression models for inheritance at Time 2

Inheritance Time 2:			
		Beta	Sig.
Model 1: AdjR^2=0.365 (p<0.001)			
	Age	0.024	0.900
	BPVS	-0.007	0.974
	Digit recall	0.186	0.142
	b/w digit recall	-0.119	0.403
	Expressive language	0.485	**0.041**
	Number of adults (comp)	0.004	0.968
	Number of younger children	-0.138	0.145
	English native	-0.099	0.362
	Parent edu/occ	0.115	0.258
Model 2:	AdjR^2=0.429 (p<0.001)		
	Age	0.069	0.703
	BPVS	-0.013	0.954
	Digit recall	0.134	0.270
	b/w digit recall	-0.115	0.397
	Expressive language	0.160	0.541
	Number of adults (comp)	-0.047	0.609
	Number of younger children	-0.138	0.133
	English native	-0.070	0.497
	Parent edu/occ	0.174	0.078
	Biodiversity Time 2	0.016	0.940
	Ecology Time 2	-0.108	0.517
	Evolution Time 2	0.484	**0.003**

Table 10.30. Composite hierarchical regression model for evolution at Time 2

Evolution Time 2			
		Beta	Sig.
Model 1: AdjR^2=0.575 (p<0.001)			
	Age	-0.060	0.700
	BPVS	0.066	0.723
	Digit recall	0.111	0.283
	b/w digit recall	-0.003	0.977
	Expressive language	0.729	**<0.001**
	Number of adults (comp)	0.122	0.106
	Number of younger children	-0.011	0.885
	English native	-0.070	0.427
	Parent edu/occ	-0.126	0.133
Model 2:	AdjR^2=0.704 (p<0.001)		
	Age	-0.121	0.355
	BPVS	-0.015	0.925
	Digit recall	0.071	0.416
	b/w digit recall	0.001	0.989
	Expressive language	0.282	0.133
	Number of adults (comp)	0.130	**0.047**
	Number of younger children	0.014	0.836
	English native	-0.020	0.784
	Parent edu/occ	-0.142	**0.046**
	Biodiversity Time 2	0.295	**0.044**
	Ecology Time 2	0.276	**0.019**
	Inheritance Time 2	0.251	**0.003**

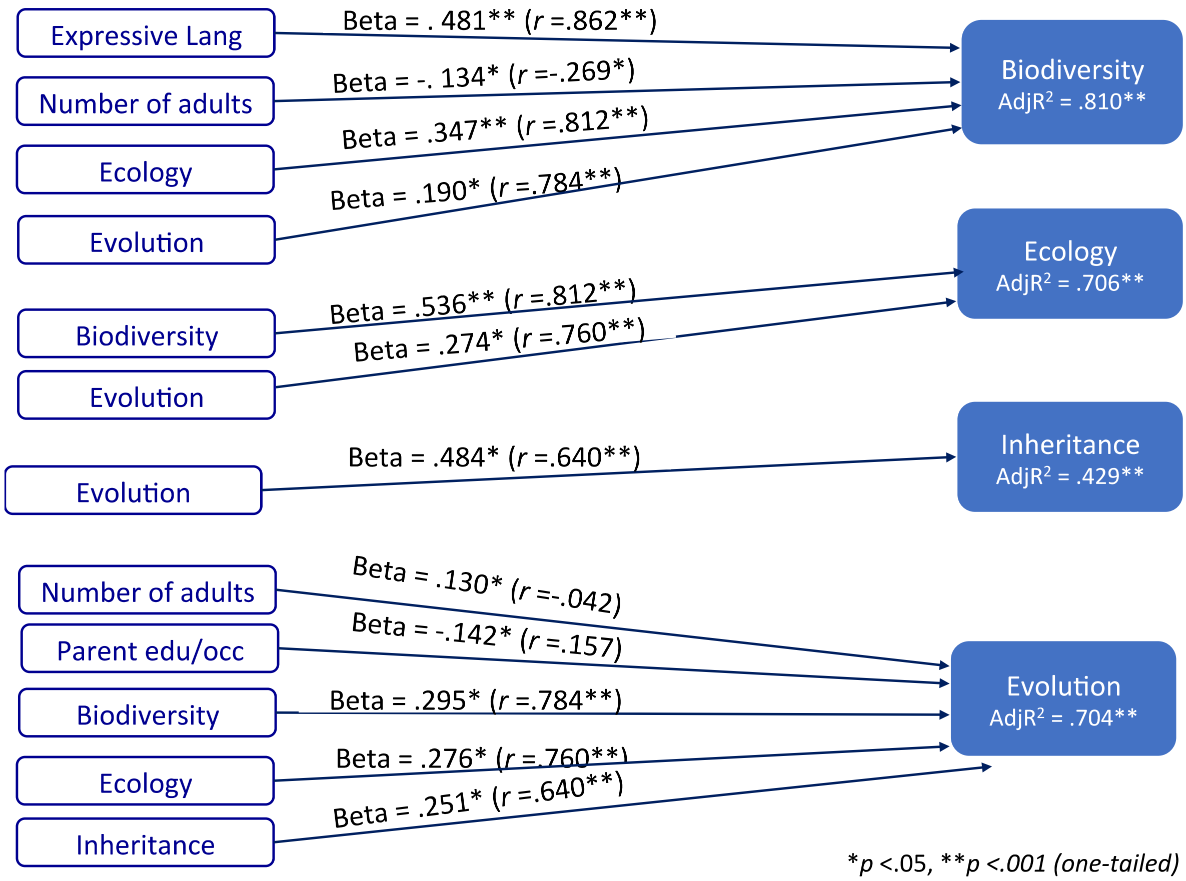


Figure 10.18. Composite hierarchical regression models for biological constructs at Time 2

10.7 Speculative path model:

Given the results of the modelling analyses above, a path model was created to try to speculate what the developmental pathway for biological knowledge appears to be. The path model was created using AMOS software. The variables chosen in the model were selected on the basis of their significance in the final composite regression models presented earlier (sections 10.6.2 and 10.6.3). The model in Figure 10.19 illustrates how much of a function language is with regards to the impact on biological knowledge.

Although this model is relatively weak (χ^2 (21) = 150.342, *p*<.0001, RMSEA= 0.212, *p*<0.001), it provides a useful illustration of how the pathways between variables appear to be. Also, the coefficients between variables are moderate to high, suggesting good relationships between them. The model suggests age, as a proxy for experience, and language ability both receptive and expressive appear to have the biggest influence on biodiversity concepts at Time 1. As suggested earlier, it appears that the specific manifestation of language ability in different aspects of biological knowledge is most important, rather than language ability per se. Also as described earlier, it seems as though expressive language is a stronger predictor of biodiversity knowledge than receptive language in this regard. These initial biodiversity concepts form a basic foundation on which ecology concepts then build at Time 1, and subsequently at Time 2. The context of ecology helps to bind piecemeal concepts together initially, and later alongside the context of evolution at Time 2. It may be that ecology and evolution allow a child to coordinate ideas and promote more divergent and abstract thinking, allowing them to develop more complex ideas about biological phenomena. This in turn influences the development of children's inheritance understanding at Time 2, although this is by far the weakest area of children's understanding.

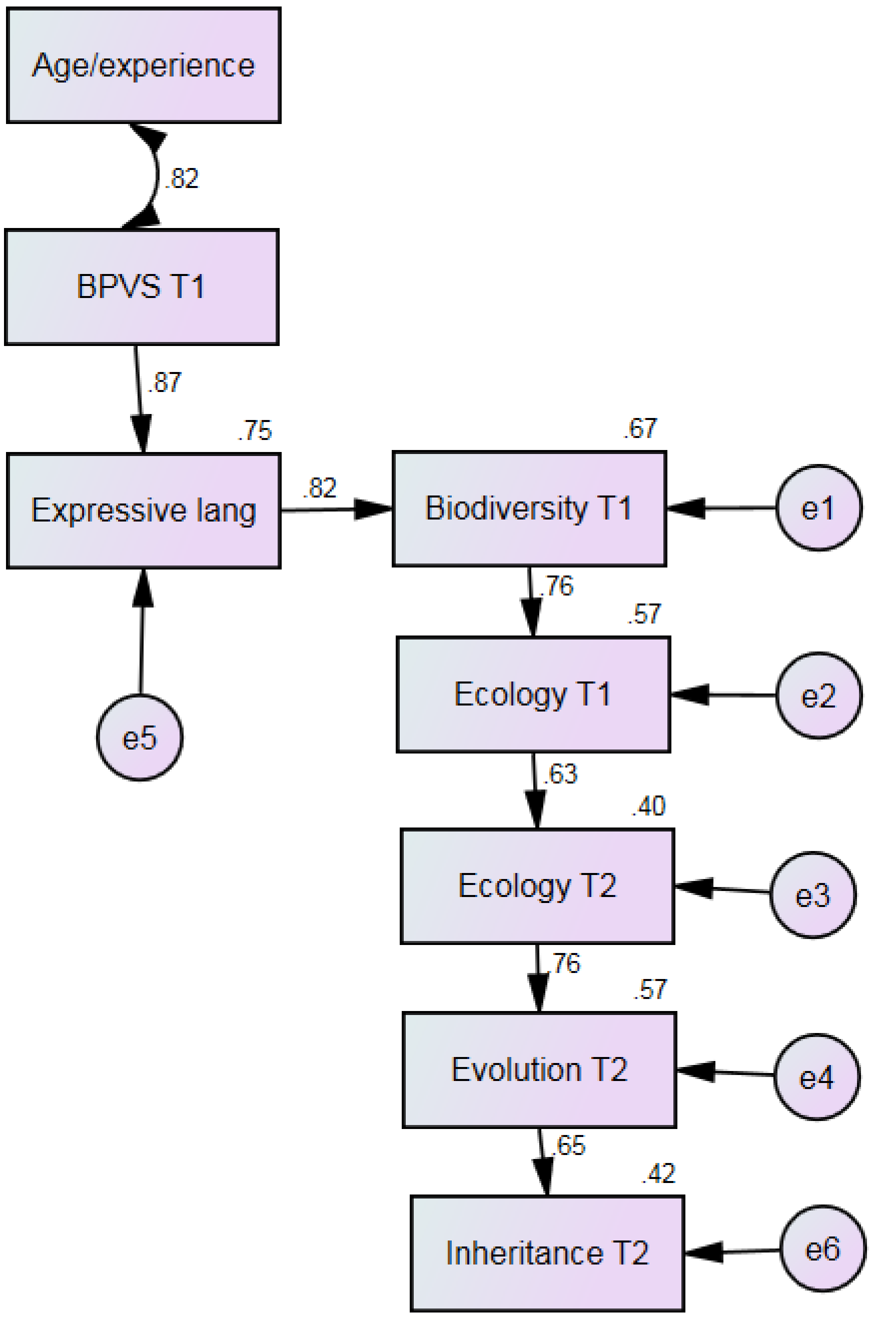


Figure 10.19. Path model of developmental trajectory of biological knowledge across the primary age-range

CHAPTER 11: THEORETICAL DISCUSSION

11.1 Overview

This chapter presents theoretical discussion of the key themes from the results presented in Chapter 10. Firstly, the poor inheritance scores at Time 1 and Time 2 are discussed as potentially resulting from irregularities, primarily in the way inheritance concepts were treated in comparison to other biological areas. One issue that was highlighted in previous analyses was that children's inheritance scores were consistently low across both time points for all cohorts. Additionally, the developmental trajectory over the two time points seemed to be cohort-specific rather than reveal any type of developmental sequence. Past research had indicated children's relative strengths in aspects of inheritance knowledge (cf. Keil, 1989; Springer, 1999; Gelman, 2003, 2015) and given that the results from the present work suggested otherwise, precautionary actions were taken to ensure inheritance concepts were not being assessed differently to any of the other biological constructs. These steps are outlined below. Elements are explored in more detail highlighting the conceptual differences between inheritance to biodiversity, ecology, and evolution. Scoring and question development of the biological constructs are additionally explored.

The results from Chapter 10 also illustrated the importance of language for children's development of biological concepts. Theoretical accounts of why this may be are discussed, as are the implications of these findings. The apparent impact of testing on learning was also demonstrated in Chapter 9. The consequence of this finding is discussed with reference to the scale of that impact relative to school influence, and what these findings suggest about the mechanisms that might be underlying conceptual development in biology.

11.2 Conceptual Analysis

One issue with inheritance was that children from all cohorts seemed to perform poorly across both time points in comparison to the other three biological constructs, and did not seem to progress over time. Also, there were very few Time 1 cognitive or demographic predictors that predicted success on inheritance knowledge at Time 2 (Chapter 10). One possibility for these results is that there is a central issue with the way questions have been devised to explore the inheritance elements taken from the NC (2001; 2014). To investigate this in more detail, a conceptual analysis was conducted investigating elements, question development, and scoring issues.

11.2.1 Elements

In total, there were 13 elements taken from the NC (DfE, 2014). These are tabulated below:

Table 11.1. Inheritance core knowledge elements.

Describe the main stages of the plant life cycle, *noting inherited traits do not change over time*
Recognise the process of growth in plants, *noting inherited traits do not change over time*
Describe how offspring of plants resemble parent plants in features
Give reasons why living things produce offspring of the same kind
Recognise that species run true
Recognise the process of reproduction as the mechanism behind inheritance

Describe how offspring of animals resemble their parents in many features
Recognise the process of growth in animals, noting inherited traits do not change over time
Describe the main stages of the animal life cycle noting inherited traits do not change over time
Describe how offspring of animals resemble their parents in many features
Describe the main stages of the human life cycle, noting inherited traits do not change over time
Recognise the process of growth in humans
Recognise and compare the main external features of the bodies of humans and other animals.

Upon closer inspection, there are eight out of the thirteen key elements that are specifically about the processes and mechanisms (highlighted in bold text) behind inheritance which inevitably posed a problem for primary school children given that they were not formally taught about the processes of sexual reproduction or genetics. The fact that only five elements out of a total of thirteen did not require explicit explanation about the processes and mechanisms behind inheritance and focused more on descriptive knowledge of facts, meant that overall children had relatively low scores for the inheritance construct on the basis of their very limited knowledge about explanatory mechanisms. The mechanisms behind inheritance posed a greater problem for children than mechanisms around the other conceptual areas because children were perhaps able to grasp and make use of existing knowledge available to them to try to logically understand and offer an explanation for the other constructs. For example, under the evolution construct children are asked "*how did the lake get to be like this*?" Children who did not know the answer made logical guesses such as "*well I'm not sure but maybe it rained a lot and there was a ditch in the ground so the water filled up like a puddle*" (child aged 9). This becomes much harder to do for something like inheritance given that it is, conceptually, very different. When children were asked a question under the inheritance construct like "*how do animals have babies*?" Those with no concept of reproduction were not able to offer any intuitive explanation and simply stated "*I don't know*."

With inheritance, there is something inherently more difficult to grasp. Many adults themselves do not fully understand the processes or genetics involved, and indeed the core aspects that enable one to understand the concept *fully and coherently* would require education beyond secondary school level; note this would also be the case for the remaining biological constructs. What must be addressed, however, is the fact that these elements are included in the NC for primary science. That is not to say it is expected that children will be able to demonstrate clear and coherent knowledge in these areas by the end of primary school, rather the markers of academic achievement would inevitably be recognised differently because children are expected to demonstrate their knowledge 'age-appropriately' as the curriculum specifies (DfE, 2014).

This study used the elements obtained from the NC in an effort to ground interview questions and does not therefore score or measure the data in the same way that a primary school teacher might, and this is worth keeping in mind. As such, the elements in the other biological constructs have been measured in exactly the same way so it is not the case that even though these elements have been included in the NC, children will automatically have equally good knowledge in all these areas.

Only five elements out of a total of thirteen elements did not require explicit explanation about the processes and mechanisms behind inheritance. The remainder of the thirteen elements would perhaps be classed as more *descriptive* or factual in nature and in this respect, the more description or the more reasons children provide for this descriptive understanding, the higher the score a child would obtain. Thus, a lack of knowledge in one particular area would lead to low scores. One point to note is that biodiversity also contains many descriptive elements (9 descriptive/13 explanatory) and yet children consistently held higher scores for biodiversity across all year groups over time. It may be that with the aid of the contextual scenes, children were able to answer more descriptive-type questions better, in comparison to inheritance questions, where cues were not available to the child in the same way e.g. there were no pictorial representations of various stages of growth in plants (in fact including such examples would have provided overt answers to the children). In this way, it could be argued that inheritance concepts were not as generously aided by the pictorial contexts as well as some of the other biological constructs. This also meant that the majority of inheritance questions were not context-specific.

The counter-argument to this is that even when children were asked generic inheritance questions, they would frequently respond with reference to the context, and therefore treated questions context-specifically. For example when children were asked "*how does a plant change as it grows*?" (Q10; element I12) children replied with answers such as "*well see like this tree here* [points to tree in the savannah] *this tree is in a hot place so it needs loooaaads of water. But it also likes a lot of sun because it doesn't like getting too cold or it won't grow leaves*" (child aged 8). This answer was typical of many children and demonstrates how they answered generic questions with reference to the contextual scene in front of them. Arguably, plants were available on the page but even in questions regarding humans, children responded similarly. When asked: "*How do animals come to have babies?*" (Q15; element I2) a 6-year-old child replied "*I think it's like this bird, the* *baby's in the tummy but we don't lay eggs.*" These examples illustrate that despite the assumed disadvantage children might have had regarding inheritance concepts by not being assisted with the contextual scene, children still frequently used the scenes as support of their answers, sometimes in an anthropomorphic manner. Inheritance questions are shown in Table 11.2 below. Note questions 4, 5, 7, and 8 all use cues from the context. The remainder of the questions are all asked in generic and non-observable terms.

It may be that because a contextual scene was placed in front of children throughout their interview and that many of the other interview questions were context-specific, children attributed salience to the scene and used it whenever possible when responding to inheritance questions. Thus, it can be argued that children treated generic and context-specific questions in the same way leading to relatively equal treatment of all questions in the interview. This is further supported by the fact that there were no contextual differences in performance or order effects for inheritance concepts (Chapters 8 and 9). This suggests that the questions targeted similar (if not the same) aspects of conceptual knowledge. Further evidence for this comes from looking at the figures showing children's responses to inheritance elements across cohorts and across contexts (see elements graphs in appendices-A.8). The graphs display a very similar pattern of consistency across both pond and savannah contexts.

The element children found the hardest was I6: *Describe how offspring of plants resemble parents in features*, responses for this were very low and seemingly lower for the savannah context, although both contexts were consistent. Overall the patterns illustrate that the highest scored items, although still scoring low, were those that were more descriptive in nature. These elements showed some growth from Time 1 to Time 2, although this was cohort-specific and scores seemed to flat-line across Cohort 2 over time points.

11.2.2 Question development

Another potential issue with the inheritance concepts is that the questions developed for the biological interview task, which were aimed to address each element for the construct, were flawed and/or not addressing the element in the intended way. As mentioned above, the inheritance elements themselves often dealt with ideas that were not as observable to children and not aided by the context they saw in front of them, but also by a (likely) lack of experiential knowledge. In some sense the questions targeting the inheritance elements were less specific than the questions for the other biological constructs simply because they had to be. Interestingly, many of these questions were human-focused in a way that the questions or elements for the other biological constructs were not. Given research into anthropomorphism in naïve biology (e.g. Inagaki & Hatano, 2004; Carey, 1985), it would have been predicted that children would actually find these questions easier given that they would be able to refer to themselves as exemplars. However, it is evident that this was not the case and the general style in which the inheritance questions were asked, more explanatory and generic as discussed above, may have posed a problem.

In the majority of the cases for inheritance elements, children were asked questions that were directly related and worded in the same manner to target a specific element (Table 11.2).

Table 11.2. Analysing inheritance core knowledge elements by interview questions asked. Yellow shading indicates no direct mapping, green shading indicates indirect mapping

1	How does a plant change as it grows?	Describe the main stages of the plant life cycle, *noting inherited traits do not change over time*
2	As above.	Recognise the **process** of growth in plants, *noting inherited traits do not change over time*
3	Do plants look the same? Why/why not? How?	Describe **how** offspring of plants resemble parent plants in features
4	(*refer to animal on the page*) Why doesn't it ever look like another kind of animal? Why does it always look like the same kind of animal?	Give **reasons why** living things produce offspring of the same kind
5	Can a (*select animal on the context*) give birth to another animal like (*select a different animal on the context*)? Why/why not?	Recognise that species run true
6	How do animals come to have babies?	Recognise the **process** of reproduction as the mechanism behind inheritance
7	If this (*select animal off context*) had a baby, will this baby look exactly like its mum and dad? More like its own mum and dad from its family, a different mum and dad from a different family, or will they all look the same? Why? How?	Describe **how** offspring of animals resemble their parents in many features
8	How will this baby (*select animal off context*) change as it grows up?	Recognise the **process** of growth in animals, noting inherited traits do not change over time
9	As above.	Describe the main stages of the animal life cycle noting inherited traits do not change over time
10	Do people ever look like their mum and dad? Do people ever look exactly the same as their mum or dad? Why/why not? How?	Describe **how** offspring of animals resemble their parents in many features

11	How does a baby change as it grows up?	Describe the main stages of the human life cycle, noting inherited traits do not change over time
12	As above.	Recognise the **process** of growth in humans
13	Do animals and people grow in the same way? What is the same/different?	Recognise and compare the main external features of the bodies of humans and other animals.

Questions not shaded on Table 11.2 ask children to produce a response that is *directly* related to the element behind that very question. This is evident in all questions except those highlighted in green, which access answers related to the element after further probing of the initial answers children give, hence these questions obtain the answer relating to the target element indirectly. Out of all the questions for inheritance displayed on the table, it is only question 1 that does not directly relate to the element, however whenever children produced answers that were off target, the researcher would re-direct them, rephrase, or probe their answers further with additional questions. It is certainly the case that questions with a direct and obvious mapping to elements are likely to specifically target children's understanding about that particular element. Similarly, even those questions shaded in green that provide a more indirect mapping would still allow examination of the element in question. However, it is possible that those questions where there is an indirect element-question mapping, may not trigger the aspects of knowledge in children that are being assessed, ultimately leading to lower scores for those particular elements.

In sum, the objection that inheritance concepts were assessed differently to the other three biological constructs by poor element-to-question mapping has been rejected. It simply appears as though the overall rate of progress for inheritance concepts across cohorts is very little. Where there is progress, it generally seems to be for more descriptive type elements and not explanatory ones. Rates of progress across Time 1 and Time 2 are also cohort-specific, with a unique pattern for Cohort 2 in that very little progress is seen here across the two time points.

In order to make sure that inheritance is not strictly unique in the way that elements are organised into more descriptive or explanatory groupings and that the questions targeted the elements in the same way for the other biological constructs as they did for inheritance, question-element mappings were also explored for the remaining constructs.

11.2.3 Biodiversity

For biodiversity, there are 15 core knowledge elements in total which are tabulated in Table 11.3:

Table 11.3. Analysing biodiversity core knowledge elements by interview questions asked. Green shading indicates indirect question-element mappings

1	Can all animals live here? Why/why not?	Recognise different animals are found in different environments, focusing on the range of organisms supported by a habitat
2	Could this animals live somewhere else? What kind of place could it live in? Probe. Can they live somewhere	Identify similarities and differences between

	icy/dry…why/why not?	environments and the effect this has on the animals that live there, focusing on the range of organisms supported by a habitat
3	What does an animal need to live? Does it get here? Can it go somewhere else to get it? Probe.	Describe **how** different habitats provide for the basic needs of different kinds of animals and **how** they depend on each other focusing on the range of organisms supported by a habitat
4	Can all plants live here? Why/why not?	Identify similarities and differences between environments and the **effect** this has on the plants that live there, focusing on the range of organisms supported by a habitat
5	What does a plant need to grow? Does it get it here?	Describe **how** different habitats provide for the basic needs of different kinds of plants [and **how** they depend on each other]
6	Can they grow somewhere else like (somewhere icy/cold/dry)? Why/why not?	Recognise different plants are found in different environments focusing on the range of organisms supported by a habitat
7	Do plants look the same? Why/why not?	Recognise all plants show variation within the same species
8	As above	Recognise all plants show variation between different species
9	Do all animals look the same? Why/why not?	Recognise all animals show variation among different species
10	Does a baby (select animal from context) always look like the mum and dad (same animal as above)? Why/why not?	Recognise all animals show variation within the same species

11	[If zebras/fish are all a bit different and have different stripes *(show child the pictorial cues*)] then how come we call them *all* zebras/fish?	Grouping living things according to their similarities and differences
12	What makes a human human? How are we different to other animals?	Recognise similarities between themselves and others
13	Do people ever look the same or do they look different to each other? What makes people look different/same?	Recognise all humans show variation within a species
14	Would it make a difference to the _if the _wasn't there anymore? Vice versa. Probe.	Understand the **process** of a feeding chain
15	You see this heron? Is it more like the trout, the frog, or the fish? Why? How is it similar? How is it different?	Identify a number of things that can be grouped as producers, consumers, predators, prey, herbivores, carnivores, etc.

Out of the 15 core knowledge elements, it could be argued that only 4 overtly ask children to explain processes or mechanisms (highlighted in bold text) whilst the remaining elements are simply about recognition or identifying facts without the need to display explicit understanding. However, unlike inheritance concepts, biodiversity concepts about variation, taxonomy, and habitats, are all visible in everyday life as well as in the pictorial contexts provided to the children. Observing regular patterns in the environment is something children have been known to be very good at from a very young age (cf. Mareschal & Quinn, 2001) thus perhaps children are likely to score relatively higher for biodiversity concepts in comparison to inheritance concepts.

In terms of the questions for biodiversity, there are 5 (highlighted in green) that indirectly target the element in question through probing initial answers. The remainder of the questions target the elements directly. Unlike inheritance, there are no questions which may potentially allude to something other than the target element. This may also explain why children have the best biodiversity knowledge across time points. The breakdown of scoring per element also shows children generally increase their knowledge across all areas aside from elements B25 and B27 where their knowledge actually takes a decrease across cohorts and contexts. It is unlikely that scoring for these elements was an issue at Time 2, because if this was the case, scoring would have flat-lined. Interestingly, these two elements are more descriptive in nature, hence reasons for the lower scores in these two elements may be due to issues around boredom. As with inheritance, it seems that elements across biodiversity are generally consistent across cohorts and contexts (appendices-A.8).

The fact that nearly all biodiversity elements show children scored similarly across items suggests that the explanatory nature did not pose too much of a problem here. In fact the lowest elements that children performed on were element B24 (testing within-species variation and the processes behind variation) and element B17 (testing interdependence of environment and organisms). It could be argued that the former element also has some crossover with inheritance knowledge, which might explain the low scores given that children struggle with this concept. The latter element overlaps with ecological knowledge and is a more explanatory variable too, suggesting children find it hard to think about issues in a more global fashion, and that systems thinking (Hipkins et al, 2008) does not really initiate until Cohort 3. Nonetheless, what is clear is that there is growth in nearly all the elements consistent across cohorts and context, implying reliability in the measures.

11.2.4 Ecology

There were a total of 10 core knowledge elements for ecology. This construct had the fewest number of elements in total, which are tabulated below in Table 11.4:

Table 11.4. Analysing ecology core knowledge elements by interview questions asked. Green shading indicate indirect element-question mapping; yellow shading indicates no direct mapping.

1	Do you think these animals live here? Why/why not?	Identify that animals live in habitats to which they are particularly suited
2	Can all animals live here? Why/why not?	Recognise different animals are found in different environments, focusing on the **interdependence** between habitat and organisms.
3	Can all plants live here? Why/why not?	Recognise different plants are found in different environments, focusing on the **interdependence** between habitat and organisms
4	As above.	Identify that plants live in habitats to which they are particularly suited
5	(*Select animal*) eats lots of grass/little water insect, if there were lots of (*same animal*) around here, would it make a difference to the (*same animal?)* Probe.	Recognise all living things are **interdependent** interacting with each other and their environment
6	If lions eat zebras, how come there are still so many zebras around in the savannah?	Recognise all living things show variation and are interdependent interacting with each other and their environment
7	Suppose it didn't rain for a long time, and the lake/pond dried up. What would happen to all the animals? Probe.	Identify the similarities and differences of different environments and how these affect the kinds of animals that live there, focusing on the **interdependence**

		between various habitats and organisms
8	[related to above question] Where would they go? What would they eat? Why?	Describe **how** different habitats provide for the basic needs of different kinds of animals and plants, and how they depend on each other, focusing on the interdependence between habitat and organisms
9	Suppose this lake dried up, would this plant be able to grow here? What would happen to it? Why?	Identify similarities and differences between local environments and how these have an effect on the plants that live there focusing on the interdependence between habitat and organisms
10	As above.	Describe how different habitats provide basic needs for plants

The number of elements that could be considered more explanatory in nature is equal to the number of elements that may be considered descriptive. There are a couple of questions (highlighted in green) above that indirectly address the elements and the child's ability to think in an ecological way, and two questions that may seem as though they do not directly target the element (highlighted in yellow). This may be because a number of ecology elements require systems thinking in that the environment and its organisms ought to be considered as a whole. This would inevitably need to be explained by the child, who would have to understand the interrelationships and explain them explicitly to answer some of the questions. Further support comes from figures for ecological elements (appendices-A.8). The elements with the lowest scores are elements Ec34 and Ec35, those which require the children to demonstrate systems thinking. These elements generally show little growth or decline in knowledge at Time 2 across cohorts and contexts. Elements Ec36 and Ec37 also show similar patterns where there is little growth between time points for Cohorts 1 and 2, but a decline in knowledge at Cohort 3 for the pond context and savannah context.

In sum, it is evident that there is growth in the majority of ecological elements which is consistent across contexts. Another interesting point to highlight is that the four elements children seemingly perform the lowest on are all explanatory-type questions, with the exception of element Ec34. This is consistent with the idea that children struggle more with explanatory type questions as they require the child to refer to processes, mechanisms and more explicit knowledge that the child may not possess. In line with this, the elements that children consistently score the highest on are Ec31, Ec32, and Ec33, which incidentally are all descriptive-type questions further supporting this idea.

11.2.5 Evolution

Finally for evolution, there are eleven core knowledge elements that are shown in Table 11.5.

Table 11.5 Analysing evolution core knowledge elements by interview questions asked. Indirect element-question mappings are shaded in green. Yellow shading indicates no direct mapping.

1	Why can't this lion/otter live in cold weather? But a polar bear can? Why? How>	Consider **how** some animals are adapted to the extreme
2	Why can't this tree grow in cold weather but other trees, like a Christmas tree can? Why? How?	Consider **how** some plants are adapted to the extreme

3	As above	Recognising **how** plants are suited to the environment in which they live
4	Do all plants look the same? Why/why not?	Recognise **why** plants produce offspring of the same kind and **link** this back to inheritance
5	Why doesn't it ever look like another kind of animals? Why doesn't it always look like the same kind of animal? (select specific animal)	Recognise **why** animals produce offspring of the same kind and **link** this back to inheritance
6	How did the Savannah/pond come to be like this?	Identify the similarities and differences of local environments and **how** these affect the kinds of animals that live there [thereby shaping the environment]
7	Do people look the same or do they look different to each other? What makes people look the same/different?	Recognise **why** humans produce offspring of the same kind and **link** this back to inheritance
8	Lions/otters eat zebras/fish, but zebras/fish are good at hiding/swimming so that lions/otters can't catch them. The zebra/fish that have lots of stripes/bigger tails are faster/better at hiding. Which zebra/fish do you think the lion/otter is likely to eat first? Why? Probe.	Recognise natural selection as the process behind evolution
9	As above	Recognising **how** animals are suited to the environment in which they live
10	As above	Identify similarities and differences between local environment and **how** these have an effect on the animals that live there
11	Suppose this lake dried up, would it better for this (choose animal) to have no baby, 1 one baby, or lots of babies? Why is that better? Probe.	Recognise natural selection as the process behind evolution

For evolution, nine out of the total eleven elements concern the processes and mechanisms behind evolution (bold text). The only example where it seems as though the question does not target the element is for question 11. It was felt necessary to directly assess for children's knowledge about natural section without actually using those terms. In this vein, question 11 was devised, and looking at the figures for element scores (appendices-A.8), children displayed quite good knowledge suggesting the element-question link is sound.

The questions that children did seem to struggle with however, were elements Ev38 and Ev41. The former concerns a very evolutionary element in that it asks children to think about *how* animals became suited to the environment in which they live *over time*. This type of 'process' question inevitably posed some difficulty. The latter recognises the process of inheritance and so once again, children may have struggled answering this question for reasons explained earlier. Generally for all other elements related to evolution, growth across time point is evident for all cohorts and this is consistent across contexts. The lowest performing elements are: Ev45, Ev43, Ev44, Ev40, Ev41, and Ev38. Upon closer inspection, all of these elements are all explanatory in nature suggesting children find explicating processes or mechanisms harder. This is to be expected however, given children's lack of formal instruction in the area. Overall the analysis and element-specific figures (appendices-A.8) indicate sound question-element mapping. Children seemed to find explanatory questions harder across all constructs and as inheritance has the most of these types of questions, children inevitably find these concepts harder to grasp.

11.3 Elements used in multiple questions

Another aspect worth exploring with regards to the equal treatment of biological constructs, is how far it is the case that the same question is addressing multiple elements, and whether there is some bias regarding inheritance in this way. The questions in the interview that target multiple and overlapping concepts are tabulated below in Table 11.6.

Table 11.6. Core knowledge elements that were used multiple times for interview questions. Eleven in total.

Question	Element	
2	B18	Recognise different animals are found in different environments - focusing on the *range* of organisms supported by a habitat
	Ec30	Recognise different animals are found in different environments - focusing on the *interdependence* between habitat and organisms
6	Ec29	Recognise different plants are found in different environments - focusing on the interdependence between habitat and organisms
	Ec32	Identify that plants live in habitats to which they are particularly suited
	B17	Identify similarities and differences between different environments and the effect this has on the plants that live there, focusing on the range of organisms supported by a habitat
9	Ev45	Consider how some plants and animals are adapted to the extreme
	Ev40	Recognising *how* plants are suited to the environment in which they live
10	I12	Describe the main stages of the plant life cycle- noting inherited traits do not change over time.
	I9	Recognise the process of growth in plants
11	I6	Describe how offspring of plants resemble parent plants in features
	Ev42	Recognise why plants produce offspring of the same kind and link this back to inheritance
	B22	Recognise all plants show variation within the same species
	B21	Recognise all plants show variation among different species

13	I3 Ev43	Give reasons why living things produce offspring of the same kind Recognise why animals produce offspring of the same kind and link this back to inheritance
18	I7 I11	Recognise the process of growth in animals Describe the main stages of the animal life cycle *noting inherited traits do not change over time*
22	B25 Ev44	Recognise all humans show variation within a species Recognise why humans produce offspring of the same kind and link this back to inheritance
24	I10 I8	Describe the main stages of the human life cycle, *noting inherited traits do not change over time* Recognise the process of growth in humans
30	Ev47 Ev41 Ev38	Recognise natural selection as the process of evolution Recognising *how* animals are suited to the environment in which they live Identify similarities and differences between local environment and how these have an effect on the animals that live there
33	Ec37 Ec34	Identify similarities and differences between local environment and how these have an effect on the plants that live there focusing on the interdependence between habitat and organisms Describe how diff habitats provide basic needs for plants…

Out of the entire interview schedule only two elements are used twice. These are Ec31, and Ev47. However both of these elements are directly targeted to specific questions and so are unlikely to pose a problem. With regards to Table 11.6 above, there are 5 biodiversity elements, 5 ecology elements, 7 inheritance elements, and 9 evolution elements. These numbers are not far from equal, aside from evolution. When one looks more closely at the evolutionary elements involved, they all consider aspects of processes, or mechanistic understanding which is why they were used alongside other related elements, such as those around ideas of inheritance because these concepts are inherently linked and (assuming good knowledge) one is likely to have to refer to one alongside the other e.g. question 18.

By this same logic, it is possible that evolutionary concepts will be present alongside other elements the most because of the very nature of its structure; further supported by the fact that there exist 8/10 explanatory elements in comparison to just 2/10 descriptive elements. What Table 11.6 also highlights however, is that there does not seem to be any bias towards/away from inheritance elements in the way that elements have been used in multiple questions.

11.4 Scoring of inheritance

An additional thing to consider with regards to children's poor scores in inheritance is the scoring system itself, and whether it biases against inheritance elements in any way. As discussed earlier, the scoring system was a scale measure; the more knowledgeable a child was about a particular phenomenon, the higher the score they received. The scoring system was designed to capture the shift in children's explanation from non-scientific, to scientific explanations. It also accounted for the child's use of logic behind a question even if that logic was inaccurate. If the explanation alluded to a scientific explanation of sorts, the child was able to obtain a relatively high score, as illustrated by the examples below:

> "*They are* borned *different*"
>
> *"because if I look whole, like all of it like mum, I'd be just like her. If I looked all of it like my dad, I'd look like him. So if it's half and half, I'm him and her."*

Thus with regards to inheritance, even if children obtained lower scores due to a lack of knowledge, if they were able to provide an answer that displayed some scientific basis it would be the case that they should obtain a relatively high mark. In this sense the scoring system could not have been the sole reason behind children's low scores in this area. Rather it must be something which did not lead children to think creatively or logically about a possible answer. It could be that the pictorial representations simply did not provide children the cues they needed for the questions However, children referred to contexts even when generic questions were asked suggesting generic questions did not impact children's answers differently between contexts as demonstrated earlier. Perhaps it is the general style of the questions, being more explanatory rather than descriptive in nature, that had more of an impact for children that meant they may not have been able to refer to their existing knowledge, which they could have been able to do for the other biological constructs (section 11.2.2). As explained above however, there were reasons to believe children would have found these types of questions easier. Also, the nature of some of the inheritance questions meant that providing cues was simply not possible.

This study made a wider attempt at examining conceptual development in biology than other previous studies have taken. The conceptual analysis above does however, dismiss any obvious artifactual results in the data. Thus, regardless of whether inheritance is conceptually different from the other three biological constructs, the evidence from the conceptual analyses provided above suggests that there were no differences in question development or scoring of these constructs. The reason for children's very poor inheritance scores are most likely due to their lack of any intuitive knowledge in this area and not methodological differences in the treatment of the biological constructs.

Likewise, with regards to evolutionary concepts there may be specific items where children display more knowledge about than others, a sub-group of micro-evolutionary concepts so to speak. This is because despite children having a problem with change over time, they are still capable of understanding some elements of particular phenomena and so may gain low scores even when overall their knowledge is incorrect, whereas for inheritance concepts if a child did not know an answer to a question they would get a score of ‘0’. For example, in response to the question *“…which zebras are likely to be eaten up first: the ones with many stripes, or one with fewer stripes? Why?”* children would generally pick the zebra with fewer stripes, which is of course correct and so they immediately get a score of 1. However when the reasoning behind this is probed, some children may give the correct answer referring to a lack of camouflage/natural selection, others would come up with an incorrect albeit fairly logical answer of something such as “*the zebra with fewer stripes is younger so it can’t run away as fast.”* Although the latter is incorrect, the child was able to answer the initial question correctly on the basis of some sound (albeit inaccurate) logic implying a partially explicit concept (Karmiloff-Smith, 1992). In contrast, the same type of naïve thinking is not possible for inheritance or ecology constructs given the nature of the questions and so if children did not know the answer they scored ‘0’.

In conclusion, thorough examination of the elements, question development, and scoring for each biological construct suggests that inheritance was not treated differently. Although it is a conceptually different type of construct with different types of elements of knowledge that children need to demonstrate, the methodology was a sound first attempt at capturing the sub-groups of concepts within all constructs and developing questions on that basis. Children’s poor inheritance scores are therefore not a result of the methodology but provide evidence of poor conceptual understanding across primary school. This finding is in contrast to past research where results may have been an artefact of the essentialist paradigm, as discussed in Chapter 2.

11.5 A sequence of acquisition

The results Chapters 8-10 outline a potential sequence of conceptual development in biology. These findings support the view that children's natural tendencies to tacitly acquire perceptual information from the environment and categorise this information (cf. Schulz et al., 2007; Csibra & Shamsudheen, 2015), allows them to subsequently gain an understanding about taxonomy, a key area of biodiversity knowledge. As outlined earlier, biodiversity is a construct that seems to be assembled by accumulating fragmented facts rather than understanding about processes or causal mechanisms (Grotza & Basca, 2003). This makes it a relatively straightforward area for children to grasp even from as young as age 4, because there is no conceptual *development* occurring as yet. However, the consequence of accumulating large amounts of fragmented knowledge is that these concepts require coordination. Past research indicated that general cognitive ability, particularly working memory may therefore be predictive of conceptual knowledge in this regard (Alloway et al., 2008), although others had not always found evidence of this (Hecht, 2002; Hecht et al., 2001).

The models presented in Chapter 10 indicated that general cognitive ability had very little effect on conceptual progression (discussed later). Rather, the regression models presented in Chapter 10 illustrate that constructs of ecology and evolution allow this coordination to occur by assisting children to use their previously acquired biodiversity concepts, and merge them with related ecological/evolutionary concepts. These findings also support the view that systems thinking (Hipkins et al, 2008) is a necessary function with which to organise knowledge and this allows the child to view concepts globally (ecology) and longitudinally (evolution), which subsequently aids them to understand biological processes and mechanisms. It may be that domain-specific conceptual progression requires having context-specific knowledge as this is what ultimately aids understanding about causal processes as Grotza and Basca show (2003) and may be why the temporal dimension of ecological and evolutionary knowledge because it is easier to understand *in situ*, as opposed to in arbitrary contexts (e.g. Maurice-Neville & Montangero, 1992).

Inheritance on the other hand is conceptually different. Children's fragmented biodiversity knowledge does not directly relate to inheritance concepts, and cannot be coordinated in any mental model in the same way, although it would appear that children try to do this given the fact that ecology and evolutionary concepts predict inheritance knowledge. This finding supports the view that in order to examine conceptual development across biological concepts, one needs to study multiple related concepts, rather than inheritance concepts alone, as others have also suggested (Williams, 2012). This is not only informative about the content of children's knowledge, but also how children integrate pieces of related knowledge to reach a sufficient level of understanding about a particular concept. The fact that ecology and evolutionary concepts predict inheritance knowledge across time points, suggests that these two constructs (ecology and evolution) are important for coordination of fragmented concepts. Indeed work by Hipkins and colleagues (2008) argued something similar and the authors stated that concepts such as ecology allow children to engage in coordinating related concepts resulting in a more globalised view, i.e. systems thinking. It may be that systems thinking requires a context in which to ground piecemeal ideas in order to coordinate them as recent research by Almeida and colleagues (2013) highlighted, but also past work that indicates contexts constrain children's thinking resulting in more accurate conceptions (Keil, 1996).

In some sense, it could be that children are using whatever pieces of knowledge they have at hand to try to understand a concept, like inheritance, that they are not too familiar with. In this regard, Carey's (1985) account conceptual change supports this idea. When children have limited biological knowledge, they seek alternative psychological explanations, which is why Carey argued a domain of biology arose from a domain of psychology.

However, if children require the context of ecology or evolution to coordinate fragmented ideas as they allow children to think about phenomena across a temporal and global axis, then past literature indicates that children are likely to have very context-specific ideas (Almeida et al., 2013; Williams & Smith, 2010; Hipkins et al., 2008; Assaraf & Orian, 2001), and yet this is not something that was evident in the results from this study. Potential reasons for why this might be the case are explored in Chapter 12.

Consequently, it appears as though the amount of experience and exposure a child has to biological phenomena, coupled with their burgeoning need to categorise and label organisms, processes, or other phenomena, which language ability and executive control aid, leads to greater knowledge acquisition in biodiversity initially than any other construct. Likewise the lack of any significant context differences at Time 1 would also suggest that the information children initially acquire could come from secondary sources (media, teaching, museum trips etc.) rather than actual physical exposure to particular environments. This is in line with Karmiloff-Smith's (1992) account of RR that suggests children's exposure to the environment allows them to develop innate context-specific predispostions and attentional biases to focus on linguistically relevant input and with time, build up modularised representations. This explanation also works hand-in-hand with research indicating that children are excellent at categorising (Csibra & Shamsudheen, 2015) and may even explain why: attentional biases. However, what is still unclear is whether or not the level of exposure a child has to biological phenomena leads to context-specific understanding.

Aside from related biological constructs being predictive of each other, the results also illustrated the impact of language on the developmental sequence. Chapter 3 outlined that language might be implicated in conceptual development by being a potential mechanism for children to be able to coordinate and explicate fragments of naïve or implicit biological concepts (cf. Karmiloff-Smith, 1992; Tolmie, 2012). Findings from this study support this view and provide some further detail in this regard. Results at Time 1 indicate that receptive language is a key predictor of all biological constructs (individually) in the first stage of a hierarchical regression model, but drops out in the second stage to be replaced by other biological constructs. This implies that language ability of the child is not important *per se*, rather it is the specific manifestation of language in different forms of biological knowledge.

At Time 2, a similar and stronger effect is seen for expressive language, which remains a significant predictor at stage two for biodiversity and inheritance in the preliminary cognitive models, and for biodiversity in the final cognitive models. This would suggest that expressive language ability of the child superseded receptive language ability in influencing conceptual development of biological concepts. Therefore, the results from the present work imply that acquiring specific biological language might be a key mechanism behind conceptual development in biology, possibly due to its role in making explicit former implicit concepts, and coordinating piecemeal ideas (Karmiloff-Smith, 1992; Tolmie, 2012). Chapter 3 highlighted the importance of the role language had in helping children to acquire biological knowledge and therefore also accounts for why early studies using language-sparse models (e.g. Keil, 1996; Springer, 1999; Gelman, 2015) gathered conflicting data that children's early biological knowledge is coherent and theoretical. Instead, by using a language-heavy method of interviews, this study has been able to demonstrate that language may so important towards children's conceptual progression that it acts as a driver, as speculated by Tolmie (2012).

Various models of conceptual change are also able to account for the importance of language in other ways: Piaget's account of cognitive conflict (Piaget, 1970), Vygotsky's account of the co-construction of knowledge (Vygotsky, 1978), and Tolmie's two-systems hypothesis (Tolmie, 2012).

Studies on collaborative group work have also highlighted the delayed benefits of collaboration, seemingly because language used in these settings to communicate points of view might have a subsequent priming effect, allowing children to become sensitised to any relevant information they come across *ad hoc* (Howe et al., 2005). In fact the findings from this study are in support of Howe and colleagues' view because children's understanding of biological knowledge (aside from inheritance) at Time 2 showed a consistent improvement from Time 1, where children at Time 2 in Cohort 1 were performing to a similar level to children in Cohort 2 Time 1 (Chapter 9). This consistent linear increase implies that the *process of testing alone* improved children's knowledge of biological constructs, because children were given the opportunity to reflect upon reasons for their ideas more deeply and engage in dialogue about these ideas during the interview task. Furthermore, the fact that this did not seem to be the case for inheritance concepts indicates that results are not a result of improving general cognitive abilities as others (Alloway et al., 2008; Zaitchick et al., 2013; Nayfield et al., 2013) suggested, and highlight the importance of language as a potential driver for conceptual change, as conjectured in Chapter 3.

The findings from this study are also in line with those found in Williams and Smith (2010) whereby language-heavy tasks, such as explaining the causal mechanisms behind phenotypic similarity, revealed children had poor understanding about biological concepts in comparison to language-sparse tasks that made use of children's concrete judgements. It seems reasonable to conclude, therefore, that children are able to acquire both implicit and explicit knowledge about biological facts. This is likely aided by the fact that children are excellent at detecting patterns of co-regularity in the environment (Shultz et al., 2007) and categorizing this information (Csibra & Shamshudeen, 2015). However, even though children are able to acquire fairly accurate and detailed facts about biological knowledge, they routinely fail to understand the mechanisms behind biological phenomena as has been demonstrated here, but also past research (Williams, 2012; Williams & Smith, 2010; Grotza & Basca, 2003). This tendency in children also explains why children have excellent biodiversity knowledge while they have very poor inheritance knowledge, as described earlier.

11.6 Fragmented or theoretical?

Developing a sequence of acquisition of related biological concepts across primary school was one of the aims of this study. Not only does it inform one about the level of children's knowledge in different conceptual areas but it allows comment on whether children's conceptual development appeared to be fragmented or theoretical in nature. Currently there is no consensus on this matter although, as discussed in Chapter 3, much of the evidence seemed to point toward children initially having fragments of knowledge that they eventually piece together, which may become more theoretical with increasing age (diSessa, 1993; 2004; Karmiloff-Smith, 1992; Tolmie, 2012), and based on these ideas, it was hypothesised as such at the start of this study.

Results from the modelling analyses (Chapter 10) however, indicated a slightly different picture. Every hierarchical regression model that was computed to examine the key predictors of performance in each biological construct (biodiversity, inheritance, ecology, and evolution) showed that in the second stages of the models where the remaining biological constructs were used as predictors alongside cognitive or demographic variables, the biological constructs consistently predicted performance of each other. This implies that in order to gain a good conceptual understanding about a biological construct, children's understanding about other related biological constructs were consistently the most important factors. This is a finding past research had not yet found but was speculated by some (Grotza & Basca, 2003; Williams, 2012) given how many biological concepts are heavily connected.

Research conducted on children's rudimentary understanding about inheritance has suggested that children have fairly coherent concepts by age 4 or 5 (Springer, 1996; 1999). The notion of coherent concepts in inheritance or indeed other areas of biology, inevitably relate to the ideas of conceptual change being a theoretical process. By this account children are able to coordinate ideas they acquire within a theoretical network and integrate related ideas within and across domains (e.g. Carey, 1985; Vosniadou, 2014).

Consequently, children might be thinking about biological ideas in more theoretical ways than hypothesised. Perhaps, as Vosniadou (2014) argues, there may be framework theories in place which allow children to organise related pieces of information that are context-specific. Indeed, Chapter 10 highlighted that ecology and evolution were constructs that allowed coordination of seemingly fragmented pieces of information. This element of coordination in itself would imply that fragments of knowledge would be loosely organised into related domains. In addition, the fact that biological knowledge did not appear to be context-dependent, implies that children may be utilising relevant information more generally, at least to begin with. Then, as Karmiloff-Smith's (1992) model suggests, with increasing exposure to environmental stimuli and language, children start to refine their ideas more context-specifically in an iterative cycle.

The successive development from implicit concepts to explicit ones, implies that a range of influences including family composition, pet ownership, location of residence (urban or rural) are likely to influence children's early biological concepts (Williams & Smith, 2010). Alternatively, it might be that children develop a range of implicit mini-theories for a particular biological phenomena e.g. inheritance, to understand it in different contexts, echoing the ideas of Vosniadou's framework theories (2014). These mini-theories may later become integrated into a unified framework to form one broader and perhaps more coherent theory that children use to guide them across a range of tasks (e.g. Williams & Smith, 2010).

However, while there might be some sort of 'framework' in place for a theoretical understanding, there is no one biological construct in the results obtained from this study that predicts all the other biological constructs, indeed Piaget (1972) also claimed this. The relationships we see in the hierarchical regressions could be reflecting the degree of overlap between the same elements in the biological constructs. What this would imply is that concepts are actually quite fragmented then, and the reason some coherence is seen is because children are implicitly noting this inherent relatedness between the concepts, perhaps without the use of any framework at all, as has been previously suggested (diSessa, 1993). Alternatively, it may be that in the cases where children have limited biological knowledge, they make use of any other information they have available to them (see Carey, 1985), which if true, would imply that piecemeal concepts might be organised around a degree of relatedness.

The research outlined in Chapters 3 and 4 highlights that generally all theorists agree perceptual information from the environment is children's source of initial information, which they are pre-disposed to acquire. Studies have also suggested that perceptual information acquired is generally implicit and piecemeal (Kallai & Reiner, 2010; Fugelsang & Dunbar, 2005; Howe et al, 2012; Mason & Just, 2015) and unlikely to be communicable as a result (Karmiloff-Smith, 1992). Taking this as a point of departure alongside the results from this study, a contemporary model of conceptualising the process of conceptual change in biology is offered.

11.6.1 Potential model of conceptual change in naïve biology

Using what research has already established as the point of departure in Chapter 4, it could be argued that on the basis of the results obtained here, there may be a model of learning whereby anticipatory systems exist in the brain to detect all perceptual information from environmental exposure, as Tolmie (2012) suggests. This effectively creates a tacit knowledge system of ideas which at this stage are piecemeal. This programme of work highlights the importance of language in driving forward conceptual change in biology, hence a process of coordinating atomistic biological ideas and coupling them with linguistic units seems to be the next step. This is also in line with the two-systems hypothesis described by Tolmie (2012) who suggests that the process of mapping language onto concepts may be the same as the processes outlined by Karmiloff-Smith (1992) in the RR model, whereby conceptual representations are redescribed at the same time as linguistic units, and only when both have reached the E3 level of representation can they be fused together. This fusion allows the concept to then be verbally communicable, which means the concept no longer is implicit, but has reached conscious access.

Tolmie (2012) provides further detail in this regard. He suggests that repeated support and exposure through relevant tasks and contexts aid the process of explication altogether. This explains why delayed effects in collaborative group work are seen (e.g. Howe et al., 2005), and also why results from Time 1-2 in this study showed a consistent improvement in children's biological knowledge (aside from inheritance). Tolmie (2012) advocates that explicit concepts overlay implicit ones, rather than supplant them altogether. This explains why even in adulthood, there has been evidence of the maintenance of naïve concepts (Schtulman & Valcacel, 2012).

However, rather than being a completely fragmented process (cf. diSessa, 1993), it may be that in the tacit system of Tolmie's model (Tolmie, 2012), knowledge is piecemeal but the subsequent process of explication by mapping conceptual units onto linguistic units requires that these fragmented concepts be organised in a manner that groups related ideas together, much like Vosniadou (2014) describes in the context of framework theories. Unlike Vosniadou, however, I argue that these frameworks are more mobile and dynamic in nature, allowing relevant contexts to trigger certain conceptual units to change and re-organise as required. This is evidenced by the fact that knowledge of different biological constructs can help children understand other related constructs. The continual dynamic shifts occurring in these loose frameworks also explains why children may have sophisticated ideas about some biological areas (e.g. biodiversity), and very naïve ideas about other areas (e.g. inheritance). Gradually, however with continuous explication of implicit concepts, the system may become more theoretical in nature, where frameworks may become more fixed as a result.

In sum, results from this programme of work suggest that knowledge is not quite as atomistic as previously thought. However, the two-system hypothesis (Tolmie, 2011) and the RR model (Karmiloff-Smith, 1992) are still able to account for what the process of change might look like, with the extension offered in the dynamic frameworks hypothesis. Tacit knowledge is relatively fragmented but with the aid and coordination of language, these ideas become explicated and may be organised in dynamic frameworks of related ideas that shift around during exposure to relevant contexts, eventually allowing concepts to become more theoretical. This is, of course, a hypothesis but it encapsulates both past research and the findings obtained from this programme of work.

11.7 General cognitive abilities

The results from Time 1 and Time 2 indicated a marginal influential role of general cognitive abilities on the development of biological concepts. This is in contrast to previous research which highlighted the predictive influence that cognitive abilities have on numeracy and literacy (Cragg & Gilmore, 2014; St Claire-Thomas & Gethercole, 2006; Nunes, Bryant, & Barros, 2012). More recent work had also suggested the potential influence of EFs on children's science knowledge at preschool (Nayfield et al., 2013) and to some extent at primary school level (Zaitchick et al., 2013). Reasons for why general cognitive abilities seemed to play relatively little role in development of biological concepts in this study are not clear. It may be that the measures used to assess general cognitive abilities were not sufficiently sensitive to test children across such a vast age range. This will be explored in Chapter 12.

Alternatively, the absence of general effects regarding the cognitive abilities may be because EFs influence the *initial* acquisition of concepts, and the processes of integrating and coordinating these concepts. If so, the aspects of children's biological knowledge that were assessed in this study were a *consequence* of the former integrative processes that may have been coordinated by general cognitive abilities and as such, failed to capture the role of cognitive abilities *in action*. For example, if this study were to present children with fragments of biological information and ask them to come with explanations about what they thought was happening, one would expect to see a more definitive influence of EFs and other general cognitive abilities because children would be asked to coordinate and integrate ideas. However, this study asked children about the consequence of that, where children may have already coordinated and organised their ideas. For these reasons the influence of general cognitive abilities appeared to be marginal because it may be in the more moment-to-moment processing of conceptual elements that a significant cognitive influence is observed.

The relatively weak influence of EFs on conceptual change in biology might also relate to the organisation and structure of knowledge. If learning were purely theoretical, it might be that EFs would have had a larger influence in being able to coordinate ideas (e.g. Vosniadou, 2014). In which case learning is more piecemeal than previously thought because as Mayer et al (2014) suggest, if it were theoretical, then the role of inhibitory control and cognitive flexibility might be especially salient. Given, however, that they were not shown to be particularly influential in this study, alludes to the fact that children's understanding of biological phenomena is not yet theoretical because if this were the case then children would have displayed accurate and coherent understanding of biological phenomena across cohorts, which they did not.

As described in Chapter 4, systems thinking (Hipkins et al., 2008) may also be a process that is particularly important for children to develop in order to hold multiple ideas about global, overarching concepts such as ecosystems or change over time. Indeed, previous work suggests that children seem to struggle with these types of concepts at primary school (Maurice-Neville & Montangero, 1992; Hipkins et al., 2008), hence perhaps children have not yet reached the stage at which they are able to make use of executive control which would be required in order to coordinate these types of global ideas.

There has been some recent work examining the effect of brain plasticity on IQ (Ramsden et al., 2011) which has shown that the changes in verbal IQ over adolescent years are not an artefact of measurement error, but correlate with changes in brain structure in regions associated with verbal functions. Although this work was with adolescents, it may be that a similar story is happening with regards to the development of EFs during childhood; changes in brain structure might be reflecting the effects and consequent influences we see EFs having on areas of conceptual development and academic achievement. These effects have been shown to be predictive in preschool (Nayfield et al., 2013) and secondary school (St-Claire-Thomas & Gathercole, 2006) however there have been few systematic studies suggesting the same for primary school children. Results from current work suggest that in the case of conceptual progression and change at primary school, effects of general cognitive factors are marginal at best possibly due to the same reasons described above.

11.8 Demographic variables

The influence of demographic data was not a key focus of this study, nonetheless these measures were included so that any potential impact of more environmental or social factors were not undetected. For these reasons discussion of these aspects are speculative in nature. The final composite hierarchical regression models in Chapter 10 essentially control for the majority of demographic variables, resulting in a largely cognitive model. That said, findings from the modelling analyses indicate the demographic variables that seemed most influential for children's conceptual development in biology were parent levels of education and occupation, the number of adults in the home, and speaking English as a native language.

With regards to the first parameter, there is well established research suggesting that low levels of parental education are associated with lower levels of academic achievement and IQ in childhood (Alexander, Entwisle, & Dauber, 1993; Duncan, Brooks-Gunn, & Klebanov, 1994). Interestingly, research also highlights SES may be influential in children's academic achievement (see Bradley & Corwyn, 2002 for a review) and yet the SES variable used in the modelling analyses (as measured by whether or not children received free school meals), did not suggest this to be the case. It is possible that father's level of occupation may have been a proxy for SES. If so, some marginal effects on children's biological knowledge were seen in the preliminary models. Indeed there have been studies investigating which aspects of SES most strongly relate to cognitive development with some (Mercy & Steelman, 1982; Scarr & Weinberg, 1978) suggesting maternal and paternal education levels are key factors.

This might be why SES as measured in this study was never a significant predictor in the hierarchical regression analyses; instead, parental levels of occupation and education already encompassed this.

In terms of understanding how demographic variables encourage the development of biological concepts, there is literature to suggest that mothers who worked in occupations with a variety of tasks and problem-solving opportunities provided more warmth and support to their children, who in turn manifested greater verbal competency (Parcel & Menaghan, 1990). Indeed the link between SES and verbal skills has been well documented (Hart & Risley, 1995; Mercy & Steelman, 1982, Hoff-Ginsberg, 1991). Hence, in terms of the results from the current study, parent levels of education and occupation, and speaking English as a native language all point towards developing children's verbal competency, which in turn help to explain the results from the cognitive regression models where both receptive language at Time 1 and expressive language at Time 2 are key influences of children's biological knowledge. This further supports the argument detailed earlier that acquiring biologically-specific language may be the driving mechanism behind conceptual change in biology, by facilitating the coordination and explication of implicit ideas (cf. Karmiloff-Smith, 1992; Tolmie, 2012).

The number of adults in the home was also a significant factor in many of the final regression models presented in Chapter 10. It was found that three adults in the home seemed to be the optimal number and generally, the third adult was reported as being a grandparent. There is some research regarding the important sources of information and support grandparents are able to provide to children. Many families in Britain now use grandparents as a form of childcare (Fergusson, Maughan, & Golding, 2007) and mothers who work part-time are most likely to receive grandparent help (Dench et al., 1999). Aside from benefits of childcare, the presence of an extra adult in the home supports parents of children in terms of emotional and material assistance, which in turn benefits the development of the child (Cochran & Brassard, 1979). Also, each new person in a child's environment could provide a new interactive style and content of activity. There is additional evidence of intergenerational learning where grandparents provide guidance and scaffolding, particularly of children of mixed ethic background (Kenner, Ruby, Jessel, Gregory, & Arju, 2007), hence additional adults in the home may be important sources of teaching by providing extra resources such as one-on-one learning opportunities through guided participation in learning with children (Gregory, Long, & Volk, 2004). Exactly why an additional adult after the optimal number of three does not appear to have the same benefits are unclear. The effect of the number of adults in the home after three were similar to having either one or two adults in the home. Perhaps with four or more adults in the home there is less opportunity for one-on-one time with the child and so the resources become limited compared to the level of having one or two adults in the home (Chapter 10).

11.9 Impact of testing

As briefly noted above, Figure 9.1 in Chapter 9 illustrates that there was a significant impact of testing on children's performance in the biological task; a finding that was relatively unexpected. The Time 1 to Time 2 shifts in children's knowledge about biological constructs (aside from inheritance) implies that children are gaining knowledge that is most likely a consequence of conceptual progression over the course of a year. The implication then, is that the biological task given to children at Time 1 was enough to shift their knowledge on by a whole school year. This is a relatively common finding in research around science interventions. For instance, a study into group-work intervention when learning about forces and evaporation, led to significant knowledge gains across a number of measures (Christie et al., 2004). Primary school children took part in one-hour lessons over 6-8 weeks devoted to the intervention, and their performance was assessed longitudinally. Results showed that 18 months later, children in the intervention group were still ahead of their peers at secondary school level.

Although the study by Christie and colleagues (2004) was over a longer period of time, the current body of work also supports the idea that minimal amounts of appropriate input makes a significant increase in conceptual development. It could be that participating in the study at one point allows children to discuss their ideas in more depth than they might have done previously, as well as engage in dialogue which may have helped children to explicate their ideas (cf. Karmiloff-Smith, 1992). This might consequently prime children to become sensitised to subsequent relevant information they encounter (cf. Howe 2005), as described earlier. Also the task may have been more interesting for children in comparison to the types of work they would normally engage in.

It is unlikely that the reason for significant improvement between time-points could only be due to improving underlying general cognitive abilities, specifically increasing language ability. While it is almost certainly true that this would contribute to improved performance, the uniformity of the findings seen in Figure 9.1 imply something more than just increasing language ability, particularly because one would expect an improvement in children's inheritance ideas at Time 2 and this is simply not observed. The lack of consistent change across cohorts between Time 1 and Time 2 for inheritance (unlike any other biological construct) would suggest that inheritance might be a different type of construct in comparison to the other three.

Inheritance has more generic questions over context-specific questions but upon reflection, all this really indicates is that children's answers are not supported by a physical contextual scene in front of them for the generic questions. Despite this however, there is evidence that children still answered generic inheritance questions with reference to the specific context in front of them, suggesting this construct is not being tested or treated in a different way, as discussed earlier. Past research had indicated that children may have context-specific ideas about inheritance depending upon the type of task that was used to elicit their knowledge (Williams, 2012; Williams & Smith, 2010), yet this study did not support this idea as children tended to display fairly consistent, although often inaccurate or no knowledge, of inheritance (discussed in Chapter 12).

Also, the lack of order effects seen at Time 1 and Time 2 indicated that only asking questions once (as for the generic questions, which inheritance has more of) did not make a significant difference to children's performance. If it were making a difference then one would expect to see significant order effects, particularly for the inheritance construct. Given that this was not the case, one can be confident that inheritance is being tested and assessed in the same way as the other three biological constructs.

These findings leave room for speculation about the kinds of teaching practices children are exposed to. Figure 9.1 implies that 40 minutes' worth of engaging dialogue about biological phenomena appears to have more of an effect on performance over an academic year, than does one year of schooling. Exactly why this may be is somewhat unclear. It could be the relative strength of the way the biological task is framed allowing children to go beyond piecemeal surface concepts, to thinking about complex ideas in a deeper and systematic manner, which children may not get the opportunity to do under formal instruction.

If knowledge gains after minimal exposure are very high as shown by this study, and that these gains might persist over time (Christie et al., 2004), why is the impact of intervention so high relative to the usual teaching practices? In an attempt to help answer this question, teachers within each school recruited in this study were sent questionnaires. Unfortunately the poverty of returned questionnaires means that concrete interpretation cannot be offered, and the conclusions offered below must be interpreted with caution given the very small sample. Nonetheless, the secondary status of this data still provided some useful information.

Firstly, as shown by the qualitative analysis of teacher questionnaires (appendix A.5), teachers appear to allow the opportunity for children to give explanations about something they are studying. This seems like a positive teaching strategy given the results from this study indicating the importance of dialogue for explicating implicit ideas. However, this strategy might be used as a tool for assessment, and the opportunity for more theoretical thinking could be lost. Also, given the large class sizes in urban schools, children do not get this opportunity on an individual level, which might encourage them to focus their ideas and actively engage in dialogue. Group-work was also a relatively popular strategy among teachers, both in mixed-ability and ability groups. However, when teachers were asked about whether children were given the opportunity to argue their viewpoints even if they were not their own, the range of responses was from very frequently to never. This suggests some inconsistency in the kind of teaching practices employed, but largely points towards the issue that children might not be exposed to conflicting ideas to their own, which past research has demonstrated is key for cognitive development (cf. Piaget, 1970; Vosniadou, 2014; Carey, 1985).

Although the number of returned questionnaires from teachers was very few, it appears as though the strategies employed in science lessons are vast, and age-specific. The current work demonstrates a task that works across primary school, and it could be that assumptions teachers might have about what children are able to do at any given age may lead to the inconsistencies in teaching techniques. Additionally, the types of professional training teachers received also varied, which may have led to teachers focusing their attention on different aspects of teaching. Nonetheless what was clear is that training appeared to benefit teachers and many changed their practices as a consequence. This implies that with further understanding about conceptual development in biology, or indeed any other area of science, the avenue for transferring this knowledge to teachers (who are receptive of this information) is sound.

11.10 Summary

Having completed conceptual analyses on all four biological constructs, it is clear that different types of questions elicit different types of responses which some children find either easier or more difficult. However, this pattern is consistent throughout cohorts and biological constructs, suggesting that although inheritance questions are conceptually quite different to those in the other three constructs, there is no bias in the way the questions have been constructed or matched to corresponding core knowledge elements.

The sequence of acquisition of concepts in the domain of naïve biology suggested the importance of language as a potential mechanism behind conceptual change. Indeed, it was speculated that language might have an important role in explicating children's implicit ideas and sensitising children towards relevant information they may encounter (cf. Tolmie, 2011). The sequence of acquisition also highlighted that the structure and organisation of children's ideas is more theoretical than previously argued in Chapter 2. It appears children's ideas are initially very piecemeal but children's developing language allows some ideas to become partially explicit and as a result, relevant pieces of knowledge are organised together in dynamic frameworks which become more fixed as knowledge becomes overtly theoretical. Finally speculative remarks were offered around the results in relation to cognitive and demographic variables, and the impact of testing on children. These areas were never intended as the primary focus of this work, and yet their inclusion provided valuable additional information.

CHAPTER 12: GENERAL DISCUSSION

12.1 Overview

There has been a wealth of research that has aimed to try and understand children's conceptions about scientific phenomena. The general consensus has been that children do not enter formal instruction as *tabula rasa*, rather they are already equipped with naïve and intuitive ideas about the world around them. Often these naive concepts can pose a problem in the classroom when they are in conflict with what is being formally taught (West & Pines, 1984, 1986; Chi et al., 1994), leading to difficulties in learning about scientific concepts. To resolve such difficulties, children must undergo a process of conceptual change in order to develop accurate scientific ideas about the world around them (Nersessian, 2003).

This chapter will discuss the outcomes of this body of research with reference to aims and hypotheses outlined in Chapter 5, and provide a brief summary of conclusions relative to each aim. Firstly, the primary focus for this research, which was to determine the nature of conceptual progression in the domain of naïve biology, is explored. This is then followed by other aspects that were explored in this thesis such as demographic and cognitive influences on conceptual development, which were not primary foci of this study. Thus these latter sections are more speculative in nature. Subsequently, the limitations of this body of work are also reviewed followed by suggestions of future work. Finally, the contributions of the present study will be considered.

12.2 Aims

The aim of this research was to investigate conceptual change and progression in the context of children's ideas about biological phenomena. Examining biology is a fruitful area in this endeavour because the concepts naturally require a need for integration and can become more complex over time, unlike other scientific areas such as physics, where concepts may become more atomistic over the course of development.

The literature presented in Chapters 2 and 3 highlighted that despite agreement about the occurrence of conceptual change, there is vast disagreement as to the processes involved. Arguments for the knowledge-as-pieces, and the knowledge-as-theory perspectives generally suggested that children's initial ideas about scientific phenomena might be fragmented. This was one of the first aims of this research: to examine what the state of children's concepts were, and whether conceptual change was a theoretical or fragmented process.

Past literature has investigated conceptual change in the context of naïve biology, yet work has been limited to specific concepts (e.g. Keil, 1994; Carey, 1985; Gelmen, 2003), specific age groups (e.g. Zaitchick et al., 2013; Springer, 1999, Keil, 1987) and specific methodologies (e.g. Gelman, 2003, 2015; Carey, 1985; Springer, 1999; Keil & Lockhart, 1999). The aim here was to account for all these past limitations by developing a novel methodological paradigm to investigate conceptual change and integration among a number of ostensibly related biological constructs across the primary school age range.

Previous work has focused extensively on children's naïve notions of inheritance concepts in the context of biology, and by using the essentialist paradigm to investigate this, came to the conclusion that children have coherent understanding about inheritance concepts (e.g. Springer, 1999; Gelman, 2003, 2009, 2015). However, as discussed in Chapters 3 and 4, the results obtained by using the essentialist paradigm in these studies, were amenable to more simple interpretation and thus another aim of this study was to develop a new methodology, which would assess children's understanding about biological concepts in a robust and grounded manner. Specifically by looking beyond biological concepts in isolation with regards to conceptual change, and study the potential influence of general cognitive abilities on scientific understanding, which past literature has indicated are influential at preschool (Nayfield et al, 2013), primary school (Zaitchick et al., 2013), and secondary school level (Gathercole et al., 2005; Alloway et al., 2005; St Claire-Thompson & Gathercole, 2006).

Having been reminded of the aims of this study, this chapter now presents a summary of the findings in light of the research questions.

12.3 Summary of findings

12.3.1 Do children have sophisticated ideas about inheritance concepts?

One of the key findings from this body of work was that children appeared to have very little knowledge about inheritance concepts in comparison to other biological concepts. This was a striking finding given that past literature has suggested children have coherent knowledge about inheritance (e.g. Gelman, 2003; Keil, 1994; Gopnik & Wellman, 1994).

However, previous work had often employed the use of the essentialist paradigm when studying children's conceptions about inheritance. The paradigm worked on the notion that essentialism is an early precursor to genetic understanding, yet the paradigm itself was quite constraining on children's responses, rather than allowing them the opportunity for discussion around their conceptual understanding as the current task did. Children's responses to *classic* essentialist tasks (e.g. Gelman 2003; Springer, 1999; Keil, 1999) could be interpreted by their natural tendency to make probabilistic judgements and categorise data as shown in Chapter 4 (Schulz et al., 2007; Strevens, 2000; Mareschal & Quinn, 2001; Younger & Fearing, 1999). In this sense, the reason children appeared to have coherent knowledge about inheritance in past studies was because the paradigm was essentially testing their ability to detect and organise everyday perceptual data. The findings at Time 1 support this notion and suggest that in actuality, children have very little understanding about inheritance concepts. This is unsurprising given that children are not formally taught the process of reproduction in primary school, whereas they *are* taught biological processes behind other biological constructs e.g. how animals are suited to the environment in which they live.

Chapter 11 discussed the conceptual differences of inheritance in comparison to other biological areas and indicated that children's poor inheritance knowledge was not a consequence of biases in question development or scoring in this study. Rather the more robust approach in assessing inheritance away from any essentialist method allowed for closer examination of children's concepts, which found contrasting results to previous work.

Consequently, one must consider whether the essentialist paradigm is a useful framework to use when trying to investigate children's early ideas about inheritance. The results obtained from this study were able to show key conceptual differences between related areas of biology implying that future research ought to consider development in these areas separately rather than assuming the same rate of conceptual progression in each, as has been done previously (e.g. Zaitchick et al., 2013). It is certainly the case that children have a tendency toward essentialism (Gelman, 2015; Altran & Medin, 2001; Soloman, 2002) however this tendency seems likely to allude to children's natural inclination to make probabilistic judgements and categorise perceptual data. If so, studying essentialism may be a fruitless endeavour (Strevens, 2000; Schtulman & Schultz, 2008; Ghazali & Tolmie, 2014), and one must start to question what evidence we have yet to gain by continued use of this paradigm. The methodology used in this programme of work, therefore, provides a good step towards a suitable alternative approach to measure the level of children's understanding in areas of biology.

12.3.2 What is the sequence of conceptual progression?

Chapter 11 highlighted that the sequence of progression of biological concepts seems to begin with acquiring fragmented biodiversity concepts. This is enabled by children's natural inclination to observe regularities in the environment and categorise data. The accumulation of fragmented concepts becomes more coordinated to allow children understanding of biological concepts more globally through the context of ecology and also longitudinally through the context of evolution. Inheritance is the last concept to be acquired because it does not directly relate to other biological concepts. Ecology and evolution concepts predict performance in children's understanding about inheritance, but overall performance in this area remains low across primary school, as described above.

A key driver of conceptual progression appears to be language. As suggested in Chapter 11, biologically-specific language allows coordination of implicit ideas by explicating them when children are exposed to relevant contexts. Language may also help children to develop partially explicit ideas (cf. Karmiloff-Smith, 1992) which ultimately assists children to organise related concepts together in a dynamic framework.

12.2.3 What is the influence of general-cognitive abilities?

The influence of general cognitive abilities on conceptual development of biological concepts was marginal. This was a surprising finding given that past research had suggested that working memory and EFs in particular, might be predictive of children's understanding at preschool (Nayfield et al., 2013) primary school (Zaitchick et al., 2013), and secondary school (Gathercole et al., 2004; St Claire-Thomas & Gathercole, 2006). For these reasons it was hypothesised that general cognitive abilities would also have a large influence on conceptual progression in this study, however this was not the case. Chapter 10 outlined the marginal contribution of EFs on conceptual change, and as discussed in Chapter 11, this study did not capture the influence of EFs *in action*, but the end-results of its influence instead. For these reasons only traces of domain-general effects on conceptual development in areas of biology were observed.

12.3.4 Are children's concepts theoretical or fragmented?

The sequence of acquisition of biological concepts outlined earlier highlighted that related biological constructs are predictive of each other over and above any predictive values of domain-general capabilities. This implies that children organise related ideas together in frameworks (cf. Vosniadou, 2014) implying that concepts are perhaps more theoretical than previously argued in Chapter 2. However, children's knowledge was not completely accurate or coherent implying that fragments of knowledge must be acquired initially where some may become partially explicated (cf. Karmiloff-Smith, 1992) whilst others remain implicit and others become explicit. It could be that concepts that are explicit or partially explicit are organised in dynamic frameworks that re-organise and solidify on the basis of increasing explicit knowledge. As discussed earlier, language appears to be a key driver for this change implying that as children are better able to coordinate their language, concepts may become more explicit, and consequently theoretical.

12.4 Limitations

Despite the findings that have been presented, there were some limitations in the present work. Perhaps the main limitation of the present work is that the measures of EF abilities might not have been sufficiently sensitive. It has been documented that EFs begin emerging in infancy and are not fully developed until late adolescence (Best & Miller, 2010; Huizingga et al., 2006; Conklin et al., 2007).

However, there are no standard measures for EFs that can be used across this entire age-range or indeed the primary school age-range. Many of the measures for inhibitory control and cognitive flexibility are either for pre-school children aged under 4 years, or older children aged 6 years or above. It was possible to test EFs using separate measures within these age groups, however this option was not chosen because results from different measures would not have been comparable given the longitudinal nature of this study. For example, children aged 4 at Time 1 would have been tested using pre-school measures of EF, such as the dimensional change card sort task (DCCS; Zelazo, 2006), however by Time 2 this measure would have been too easy for children and they would have likely reached ceiling. For these reasons it was decided that measures aimed at older children would be most appropriate, given that by Time 2, the majority of children would be within the appropriate age brackets of these tasks.

Consequently, this meant that at Time 1, children in Cohort 1 (aged 4-5) were at risk of being at floor level in some of the tasks. Results at Time 1 for both the mean scores of WCST and the Stroop task indicate the developmental pattern that would have been predicted, in that children become better at these tasks with age. Closer inspection revealed that the differences in performance between the older two cohorts were not significantly different for both tasks however. This might imply that by around age 7/8 (Cohort 2) children are reaching their peak. Especially because results for Cohorts 2 and 3 at Time 2 revealed no significant differences in performance for the WCST, and only a significant difference between Cohorts 1 and 3 on the Stroop task. However, despite the lack of many significant findings in this area, the pattern of results *did* suggest influences of general cognitive abilities (albeit marginal) that were simply surpassed by language ability and knowledge of related biological concepts. This in itself provided useful information that other studies have not previously found.

The WCST itself was chosen as a measure of cognitive flexibility because it is one that has been used frequently in the past (Chelune & Baer, 1986; Romine, Lee, Wolfe, Homack, Geroge, & Riccio, 2004). The chimeric animals Stroop task (Wright et al., 2003) was chosen as a measure of inhibitory control because it was an effective non-verbal measure that could be computerised and was considered enjoyable for the child. A potential drawback of using this task was that the software used in the initial version was not obtainable and the task had to be altered, as described in Chapter 7. Altering the task may have meant that accuracy could have been a problem, particularly because the experimenter was in control of showing children the images as fast as possible. However, the effect of semantic inhibitory control *was* a significant predictor in some of the cognitive regression models illustrated in Chapter 10, which suggests that the chimeric animals Stroop task was sufficiently sensitive.

Furthermore, the chimeric animals Stroop task was chosen as an alternative to the go/no-go task (Fillmore & Rush, 2002), which would have required longer administration time and given that the testing period for domain-general capability measures was already very lengthy, the go/no-go task may have caused boredom or distress and potentially resulted in the child being out of class for longer periods of time.

Despite the fact that it appeared as though Cohorts 2 and 3 had reached ceiling on the WCST at Time 2, the Stroop task well worked at Time 1 and both EF measures worked well at Time 2. This implies that overall the measures for EFs were sensitive, particularly the Stoop task, and particularly at Time 1. With regards to the WCST at Time 2, although Cohorts 2 and 3 may have reached ceiling, the results of the study still allowed one to view the developmental trajectory across the two time points and still provided the valuable information that overall, EFs and indeed other domain-general abilities have little involvement in conceptual change in the context of naïve biology. Rather language seems to be a key driver.

With regards to language measures, another minor drawback of this study was that the expressive language measure was only administered at Time 2. This was because the expressive language measure could only be administered to children aged 6 and above; even at Time 2, there were a few children in Cohort 1 that were under the required age-range. An appropriate measure of expressive language was not obtainable for children aged 4 and above, and a receptive language measure was used instead for both time points. This measure demonstrated the important information that biologically-specific language might be a key driver of change in biology concepts, and that expressive language as measures at Time 2 might be even more so. In Chapter 10, when the expressive language measure was removed from the regression models, the receptive language measure replaced it as a significant predictor. This suggests that the effects of receptive and expressive language are similar, but the effect of the latter is stronger, possibly because the task measuring biological knowledge in children is also expressive in nature. Overall, while it would have been interesting to view the influence of expressive language at Time 1, the results obtained from the study still contributed vastly to the discovery that biologically-specific language is of great importance for conceptual change.

An additional limitation of this study it that the sample was recruited from across three schools. This is an issue because the teaching practices and biology topics covered in each school were likely to vary. The nature of recruiting a large sample meant that children could not be obtained from one school alone and in this case, three schools were necessary to gain a sufficiently large sample in order to conduct the planned quantitative analyses. Schools were recruited on the basis that they shared similarities such as a similar profile of intake of children from mixed ethnic backgrounds and were all state schools in deprived areas of London. However the school sizes did vary, with one school having double the class sizes of another, and this may have affected the resources available to children when learning, potentially leading to differences in children's knowledge. An attempt was made at trying to assess these types of differences by distributing questionnaires among the teachers about their levels of qualifications, training, and teaching practices. Unfortunately the rate of returned questionnaires was exceptionally low and none of the intended school comparisons could be made, instead a qualitative analysis of teachers across schools and year groups in general was conducted. This analysis suggested that overall, teachers employed similar teaching strategies and had similar levels of training but due to the small sample, these results must be interpreted with caution.

One point to consider about the nature of conceptual development of biological knowledge in young children, is the task used to examine this information. An interview based assessment meant that there was a potential risk of children giving particular responses which might not have truly reflected what they knew. There is no concrete way of establishing to what extent verbal assessment protocol impacted on children's performance, yet the alternative, a non-verbal assessment, would be equally hard to establish given the nature of the conceptual content involved. Piaget (1971, 1972) and Karmiloff-Smith (1992) argue that in order for children to demonstrate a sound understanding of a particular concept, that concept must be articulable. Other paradigms employing reduced verbal assessment via paradigms such as those described in section 3.1.3, have far more flaws and an absence of what children know in more comprehensive terms. Hence a verbal assessment is an appropriate method in this regard.

Additionally, serious attempts were made to ensure that children's true level of understanding was measured firstly by developing a coding scheme that mapped out progressive levels of cognitive advance. This allowed examination of children who may have had limited explicit knowledge but made a sensible attempt to provide a logical answer (even if that answer was incorrect), in comparison to a child who had no real understanding of a particular concept at all. Children's answers were also routinely probed to ensure that the exact level of understanding could be ascertained and coded accurately, rather than doing this on the basis of a 'surface response' because aspects of social desirability and demand characteristics may have been present.

Secondly, the nature of the interview questions were themselves grounded in the NC despite the fact that the structure of the NC itself is relatively uninformed (Chapter 1). The NC provided a standard from which to design the measure and against which to judge the outcome, rather than deciding this in a completely bottom-up fashion as discussed earlier in Chapter 6. There was no past work that addressed the broad sweep of biological understanding from a cognitive-developmental perspective and as such, there is no curriculum independent standard for measurement to calibrate against. One possible existing measure that might have arguably had this status are past Key-Stage tests (Standards and Testing Agency, 2013) however these themselves are not entirely objective because they are based on the NC, and the coverage of biology is only tested in one strand whereas this study aimed to address children's knowledge and understanding in a more detailed fashion. Therefore, by developing questions situated in the NC and providing visual props, children were not restricted in their answers and could be as abstract or as specific in their response and they wanted to or indeed, were able to.

In sum, taking these protocols were all appropriate attempts at mitigating the verbal demands of the interview task, to get as close as possible to the level and content of children's biological knowledge by developing an objective and principled scheme of conceptual change as possible.

Another point to consider is the validity of the biological interview and coding scheme. The core knowledge structures taken from the NC (DfE; 2014) were mapped on to an interview question aimed at targeting that specific element. In many cases this resulted in a direct one-to-one mapping, however in others, there was often one question used to target several core knowledge elements, or vice versa. This could have potentially meant that the subtleties of children's understanding about certain biological concepts may have been lost. Also the coding of these elements was fairly subjective. On the other hand, the reliability and validity analyses presented in Chapter 8 suggest that the task was valid and the coding reliable, with percentage agreement being very high at 86% on average. Further exploration of the biological task in Chapter 11 also highlights that this task was a successful initial attempt at a methodology that steered away from the essentialist paradigm, and provided some valuable insights on children's knowledge about biological phenomena that had not been reported before. Hence issues around validity of the biological interview and coding scheme are unsubstantiated.

Chapter 11 also debunked any potential indication that results from the study that suggested children's poor ideas about inheritance were a result of treating inheritance core knowledge structures differently to others. While it does appear to be the case that inheritance itself is conceptually different to the other biological structures included in this study, the development of the biological task was able to illustrate these differences, and as such, provides a valuable tool for further exploration of these ideas.

The use of only two contexts may also have been a potential limitation in the current work. Piloting studies had suggested that the two contexts used in this study were ones that worked well because children had the most knowledge on them. However, the fact that in the pilot studies performance on the other contextual scenes may not have been as high, may indicate that children's ideas about biological phenomena are context-specific (Hipkins et al., 2008; diSessa, 1993; Vosniadou & Ioannides, 1998). This study found very few contextual differences in children's performance possibly because the use of the two contexts effectively balanced this out. At Time 2 there was a tendency for children to demonstrate more knowledge for the pond context that was not present at Time 1. This may have been because this is a context children are more readily exposed to and perhaps the subsequent exposure allowed children to become sensitised to relevant ideas as discussed in Chapter 11. This study was required to limit testing to two contextual scenes because of the time restraints in testing children, but also in the programme of study as a whole. It may be that in future studies, a selection of contexts ought to be used to assess contextual differences in more detail. Alternatively it may be that providing children with opportunities to think along more general and abstract terms rather than be constrained by contexts *per se*, may have allowed them to source knowledge in similar ways across contexts and make use of commonalities across related knowledge structures. This would also lend support to the dynamic frameworks hypothesis (Chapter 11) and indicates that children's early biological knowledge may be more theoretical than initially thought.

Finally, it could be argued that a longitudinal study looking across the entire age range would have provided valuable insights across and within individuals. However, there were time constraints within this programme of work, and one of the ways in which to obtain developmental data across the primary age span was to adopt a triple-cohort longitudinal study. This allowed useful cross-cohort comparisons, but also allowed observation of the developmental trajectory across two time-points, providing valuable information about conceptual development in biology, which would have otherwise been untapped.

12.5 Future directions

While this research was able to fulfil the aims it set to, the results obtained from the present work also highlighted some additional questions that future work could explore. Firstly, the role of language as a mechanism for conceptual change in biology needs to be investigated in more detail. While there is work in support of the RR model (Karmiloff-Smith, 1992) the two-systems hypothesis (Tolmie, 2011), and the dynamic frameworks hypothesis proposed here have yet to be investigated. The neurological basis of differences between tacit and explicit understanding have only been observed in adults (e.g. Kallai & Reiner, 2010; Mason & Just, 2015) and while it has been suggested that two systems might also be present in children (Howe et al., 2005; Tolmie, 2012) more research into whether there are differences in implicit and explicit beliefs in children, needs to be conducted. If it is the case that there are differences, this would provide some support to the idea that children's initial tacit knowledge might be piecemeal and that theoretical understanding may develop over time. This would also highlight the need to assess conceptual change over a period of time, past the primary age range, to discover the points at which children's understanding becomes more theoretical in nature, whether there are key markers to this, and how it might be achieved faster or enhanced for the purpose of education. The current NC does highlight the use of scientific language and teachers reported that they frequently use relevant jargon in lessons. This appears to be a positive step towards improving conceptual development of scientific concepts.

Additionally, if EFs are thought to be chief in coordinating theoretical thought (cf. Opfer et al., 2012; Cragg & Gilmore, 2014; Zaitchick et al., 2013) more examination is needed to investigate this further. Perhaps with the development of more age-appropriate methods to assess EFs across the primary age-range, one will be able to observe their effects with more sensitivity than was possible in this study. Also, there is room for studies employing the use of neurological techniques to assess the role of EFs in conceptual change in action. Past research has demonstrated brain activation in frontal lobe areas associated with executive control (Alvarez & Emory, 2006), although employing this method would rule out examination of younger cohorts. It may be possible to replicate the study by Kallai and Reiner (2010) using older children and adolescents. This would allow investigation of potentially separable implicit/explicit knowledge systems, and whether there are any developmental changes associated with this.

Lastly, if it is the case that boosting executive control might boost theoretical thinking particularly in later years, then designing an intervention for children to increase their EF and domain-general abilities will be key in promoting conceptual change in biology. It may even be beneficial for conceptual change in other scientific areas, although this is something else that would also need to be investigated not least to identify whether the findings obtained from this study are specific to biology or whether they might also apply across domains. Likewise designing an intervention allowing children to explicate various scientific phenomena and consequently expose children to relevant contexts and information, would also allow examination of whether the findings from this study are specific to biological phenomena and whether a potential teaching aid could be developed to promote conceptual development in primary science lessons via the use of scientific language.

12.6 Summary & contributions of this work

The current work provided new insights about the nature of conceptual change in the domain of naïve biology. The key findings were that biologically-specific language seems to be a key driver in promoting conceptual change, and that domain-general capabilities have little influence over this change. Language seems chief to conceptual progression, yet the current work cannot offer concrete thoughts about the exact mechanisms behind this process. Theories by Karmiloff-Smith (1992) and Tolmie (2012) suggest some possibilities about the role of language in explicating naïve ideas, which implies that initial fragments of knowledge may become more theory-like over time.

This programme of work was also able to comment on the nature of children's understanding about inheritance. Past research suggests children's ideas about this concept are fairly sophisticated and coherent (Gelman, 2003; Springer, 1999; Keil, 1994) but by developing a new methodology to assess children's understanding, the opposite was found; inheritance is the weakest area of understanding in comparison to other biological constructs. The development of a more robust methodology allowed for systematic assessment of all sub-concepts within inheritance, and indeed other biological areas, resulting in more thorough study.

At the end of Chapter 5, three potential routes of conceptual progression were hypothesised: that knowledge is theoretical and inheritance concepts would develop first, that knowledge is context-specific and ideas might be somewhat isolated to begin with, but eventually start to inform each other, and finally that knowledge is piecemeal, with children piecing together information in appropriate contexts. After obtaining results, it would seem as though the second hypothesis seems most likely to be the case. Children's ideas appear to be perceptual in nature and an experiential route of progression occurs whereby biodiversity concepts arise first, given children's tendencies to detect regular patterns in the environment and categorise data, followed by ecology concepts as others (Hipkins et al., 2008) have also suggested children have a good grasp of. This is then followed by evolutionary ideas which, with the assistance of ecological concepts, feed into ideas about inheritance.

A somewhat surprising insight to result from the current work is that general cognitive abilities have seemingly very little influence on conceptual change. This was unexpected given the influence and in some cases predictive nature they have on areas on literacy and numeracy (Gathercole et al., 2003; Cragg & Gilmore, 2014; Alloway et al., 2008) and in preschool science learning (Nayfield et al., 2013). It also highlights the need to examine the effect of domain-general abilities over a significant length of time, as well illustrating that science is ontologically very different to literacy and numeracy and as such, teaching techniques and interventions would be specific to either numeracy, literacy, or science.

12.7. Conclusion

Questions into what the content of children's concepts is at different ages, how concepts develop with age, and how they are organised have puzzled curious minds for decades. Research aimed at understanding when children are able to learn to concepts, the mechanisms they employ to learn them, and how naïve concepts influence each other to develop into sophisticated ideas can ultimately inform teaching methods and curricula to enhance education as a practice.

This thesis asked many of the same questions outlined above and sought to examine conceptual development of related biological concepts in order to establish a developmental trajectory of change in primary-aged children. The results from the longitudinal study revealed children's concepts are organised into related bundles, encouraging coordinated thought with increasing age.

Children's ideas about biological phenomena seem to be acquired from early experiences with the environment. This provides some evidence in support of other work suggesting that predispositions to perceptual data forms the basis of tacit knowledge (Strevens, 2000; Karmiloff-Smith, 1992; Tolmie, 2011; diSessa, 1993; Kallai & Reiner, 2010). Tacit pieces of knowledge become explicated by aid of biologically-specific language, possibly through a process of representational redescription, which allows one to organise and consequently coordinate these ideas in more theoretical ways. The evidence for this comes from the results in Chapter 10, which demonstrate that children's understanding of previously acquired biological concepts significantly predicts their knowledge of related areas of biology.

The apparent influence of related concepts in conceptual progression was one of the key findings obtained from this work and highlights the need to consider multiple and associated concepts when studying conceptual change of any kind. Although there is still a substantial distance to go in our investigation, this study has made the distance that much shorter.

REFERENCES

Alexander, K.L., Entwisle, D.R., & Dauber, S.L. (1993). First-grade behaviour: its short- and long-term consequences for school performance. *Child Development, 64*, 801-814.

Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2008). Evaluating the validity of the Automated Working Memory Assessment. *Educational Psychology*, *28*(7), 725–734. doi:10.1080/01443410802243828

Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A.-M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology*, *87*(2), 85–106. doi:10.1016/j.jecp.2003.10.002

Alloway, T. P., Gathercole, S. E., Adams, A.-M., Willis, C., Eaglen, R., & Lamont, E. (2005). Working memory and phonological awareness as predictors of progress towards early learning goals at school entry. *British Journal of Developmental Psychology*, *23*(3), 417–426. doi:10.1348/026151005X26804

Almeida, A., Vasconcelos, C. M., Strecht-Ribeiro, O., & Torres, J. (2013). Non-anthropocentric Reasoning in Children: Its incidence when they are confronted with ecological dilemmas. *International Journal of Science Education*, *35*(2), 312–334. doi:10.1080/09500693.2011.608387

Altran, S. Medin, L. et al. (2001). Folkbiology doesn't come from folk psychology. *Journal of Cognition and Culture*, *1*(1), 3–42.

Altran, Scott; Medin, D. (2008). *The Native Mind and the Cultural Construction of Nature*. Cambridge, Massachusetts, USA: Massachusetts Institute of Technology Press.

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, *8*(2), 71–82.

Assaraf, O. B.-Z., & Orion, N. (2005). Development of system thinking skills in the context of earth system education. *Journal of Research in Science Teaching*, *42*(5), 518–560. doi:10.1002/tea.20061

Astuti, R., Solomon, G. A., & Carey, S. (2004). I. Introduction. *Monographs of the Society for Research in Child Development*, *69*(3), 1–24. doi:10.1111/j.0037-976X.2004.00297.x

Atran, S. (1987). The Essence of Folkbiology: A Reply to Randall and Hunn. *American Anthropologist*, *89*(1), 149–151. doi:10.1525/aa.1987.89.1.02a00120

Bang, M., Medin, D. L., & Atran, S. (2007). Cultural mosaics and mental models of nature. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(35), 13868–74. doi:10.1073/pnas.0706627104

Barak, M., & Dori, Y. J. (2011). Science Education in Primary Schools: Is an Animation Worth a Thousand Pictures? *Journal of Science Education and Technology*, *20*(5), 608–620. doi:10.1007/s10956-011-9315-2

Beddington, J. (2008). Final Project report. In J. Beddington (ed.), *Foresight Mental Capital and Wellbeing Project.* London: Government Office for Science.

Best, J., & Miller, P. (2010). A developmental perspective on executive function. *Child Development*, *81*(6), 1641–1660.

Bloom, P., & Weisberg, D. S. (2007). Childhood origins of adult resistance to science. *Science, 316,* 996-997.

Bonawitz, E. B., van Schijndel, T. J. P., Friel, D., & Schulz, L. (2012). Children balance theories and evidence in exploration, explanation, and learning. *Cognitive Psychology*, *64*(4), 215–34. doi:10.1016/j.cogpsych.2011.12.002

Bonawitz, E., Denison, S., Griffiths, T. L., & Gopnik, A. (2014). Probabilistic models, learning algorithms, and response variability: sampling in cognitive development. *Trends in Cognitive Sciences*, *18*(10), 1–4. doi:10.1016/j.tics.2014.06.006

Boucher, J., Pons, F., Lind, S., & Williams, D. (2007). Temporal cognition in children with autistic spectrum disorders: tests of diachronic thinking. *Journal of Autism and Developmental Disorders*, *37*(8), 1413–29. doi:10.1007/s10803-006-0285-9

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, *53*(1), 371-399.

Bramwell-Lalor, S., & Rainford, M. (2014). The Effects of Using Concept Mapping for Improving Advanced Level Biology Students' Lower- and Higher-Order Cognitive Skills. *International Journal of Science Education*, *36*(August), 839–864. doi:10.1080/09500693.2013.829255

Brandone, Amanda, C., & Gelman, S. A. (2013). Generic language use reveals domain differences in young children's expectations about animal and artefact categories. *Cognitive Development*, *28*, 63–75. doi:Brandone, Amanda, C., & Gelman, S. A. (2013). Generic language use reveals domain differences in young children's expectations about animal and artefact categories. Cognitive Development, 28, 63–75. Retrieved from Brandone, Amanda, C., & Gelman, S. A. (20

Brown, C. H. (2001). Folkbiology. *American Anthropologist*, *103*(1), 254–255. doi:10.1525/aa.2001.103.1.254

Brown, D. E. (2010). Students' Conceptions - Coherent or Fragmented? And What Diffeence Does It Make? In *National Association of Research in Science Teaching* (p. 17). Philadelphia.

Bruer, J. T. (1997). Education and the Brain: A Bridge Too Far. *Educational Researcher*, *26*(8), 4–16. doi:10.3102/0013189X026008004

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, *33*(3), 205–28. doi:10.1080/87565640801982312

Bull, R., & Scerif, G. (2001). Executive Functioning as a predictor of Children's Mathematics Ability: Inhibition, Switching, and Working Memory. *Developmental Neuropsychology*, *19*(3), 273–293.

Bynner, J., & Parsons, S. (2005). *Does Numeracy Matter More?* London: National Research and Development Centre for Adult Literacy and Numeracy, Institute of Education.

Byrne, J., Grace, M., & Hanley, P. (2009). Children's anthropomorphic and anthropocentric ideas about micro-organisms. *Journal of Biological Education*, *44*(1), 37–43. doi:10.1080/00219266.2009.9656190

Carey, S. (2000). Science Education as Conceptual Change. *Journal of Applied Developmental Psychology*, *21*(1), 13–19. doi:10.1016/S0193-3973(99)00046-5

Carey, S. (1999). Sources of conceptual change. In E. K. Scholnick, K. Nelson, S. A. Gelman, & P. H. Miller (Eds.), *Conceptual development Piaget's legacy* (pp. 293–326). Lawrence Erlbaum Associates.

Carey, S. (2011). Precis of the Origin of Concepts. *Behvioural and Brain Sciences*, *34*, 113–167. doi:10.1017/S0140525X10000919

Carey, S. (1985). *Conceptual Change in Childhood*. Massachusetts: MIT Press.

Carey, S. (1986). Cognitive science and science education, *American Psychologist, 41,* 1123-1130.

Carey, S. (1987). Theory change in childhood. In B. Inhelder, D. de Caprona, & A. Cornu-Wells (Eds.), *Piaget today* (pp. 141-163). London: Erlbaum.

Carey, S. (1988). Conceptual differences between children and adults. *Mind & Language, 3,* 167-181.

Carey, S. (1989). Conceptual differences between children and adults. *Mind and Language, 3*,167-181.

Carey, S. (1991). Knowledge acquisition: Enrichment or conceptual change? In S. Carey & R. Gelman (Eds.), The epigenesis of mind: Essays on biology and cognition (pp.257-292). Hillside, NJ: Lawrance Erlbaum Associates.

Carey, S. (1992). The origin and evolution of everyday concepts. In R. Giere (Ed.), Cognitive models of science (pp. 89-128). Minneapolis, MN: University of Minnesota Press.

Carey, S. (1999). Sources of conceptual change. In E. K. Scholnick, K. Nelson, S. A. Gelman, & P. H. Miller (Eds.), *Conceptual development: Piaget's legacy* (pp.293-326). Mahwah, NJ: Lawrence Erlbaum Associates.

Carey, S. (2000a). Science education as conceptual change. J*ournal of Applied Developmental Psychology, 21*, 13-19.

Carey, S. (2000b). The origin of concepts. *Journal of Cognition and Development, 1, 37*-41.

Carey, S., & Spelke, E. (1994). Domain-specific knowledge and conceptual change. In L. A. Hirschfeld & S. A. Gelman (Eds.). *Mapping the mind: Domain specificity in cognition and culture* (pp.169-200). Cambridge: Cambridge University Press.

Carey, S; Gelman, R. (Eds.). (1991). *The Epigenesis of Mind: Essays on Biology and Cognition*. New Jersey; USA: Lawrence Erlbaum Associates, Publishers.

Carey, S., & Spelke, E. (1996). Science and core knowledge. *Philosophy of science, 63,* 515-533.

Carvalho, P. F., & Goldstone, R. L. (2015). What you learn is more than what you see: what can sequencing effects tell us about inductive category learning? *Frontiers in Psychology*, *6*(April), 1–12. doi:10.3389/fpsyg.2015.00505

Chanen, A. M., & Kaess, M. (2012). Developmental pathways to borderline personality disorder. *Current Psychiatry Reports*, *14*(1), 45–53. doi:10.1007/s11920-011-0242-y

Chi, M. T. H., Slotta, J. D., & De Leeuw, N. (1994). From things to processes: A theory of conceptual change for learning science concepts. *Learning and Instruction*, *4*(1), 27–43. doi:doi: 10.1016/0959-4752(94)90017-5

Chi, M. T. H., Hutchinson, J. E., & Robin, A. F. (1989). How inferences about novel domain-related concepts can be constrained by structured knowledge. *Merrill-Palmer Quarterly*, *35*(1), 27–62.

Chin, C., & Teou, L.-Y. (2010). Formative assessment: Using concept cartoon, pupils' drawings, and group discussions to tackle children's ideas about biological inheritance. *Journal of Biological Education*, *44*(3), 108–115. doi:10.1080/00219266.2010.9656206

Cimpian, A., & Salomon, E. (2015). The inherence heuristic: An intuitive means of making sense of the world, and a potential precursor to psychological essentialism. *Behavioral and Brain Sciences*, *37*(05), 461–480. doi:10.1017/S0140525X13002197

Cochran, M. M., & Brassard, J.A. (1979). Child development and personal social networks. *Child Development, 50(3),* 601-616.

Coley, J. D. (1995). Emerging Differentiation of Folkbiology and Folkpsychology: Attributions of Biological and Psychological Properties to Living Things. *Child Development*, *66*(6), 1856–1874. doi:10.1111/j.1467-8624.1995.tb00970.x

Coley, J. D. (2000). On the Importance of Comparative Research: The Case of Folkbiology. *Child Development*, *71*(1), 82–90. doi:10.1111/1467-8624.00121

Coll, R. K., France, B., & Taylor, I. (2005). The role of models/and analogies in science education: implications from research. *International Journal of Science Education*, *27*(2), 183–198. doi:10.1080/0950069042000276712

Cook, C., Goodman, N. D., & Schulz, L. E. (2011). Where science starts: spontaneous experiments in pre-schoolers' exploratory play. *Cognition*, *120*(3), 341–9. doi:10.1016/j.cognition.2011.03.003

Council for Environmental Education. (1997). *Educating for Life: Guidelines for Biodiversity Education*. (E. McLeish, Ed.). Reading, UK: Hobsons Academic Relations.

Cragg, L., & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. *Trends in Neuroscience and Education*, *3*(2), 63–68. doi:10.1016/j.tine.2013.12.001

Chelune, G.J., & Baer, R.A. (1986). Developmental norms for the Wisconsin card sorting test. *Journal of clinical and Experimental Neuropsychology, 8(3),* 219-228.

Christie, D., Tolmie, A., Howe, C., Topping, K., Thurston, A., Jessiman, E., Livingston., & Donaldson, C. (2004). *The impact of collaborative group work in primary classrooms and the effects of class composition in urban and rural schools.* Teaching and Learning Research Programme annual conference, Cardiff, UK.

Csibra, G., & Shamsudheen, R. (2015). *The role of ostensive demonstration and object labelling in promoting generalization of non-verbally demonstrated object properties.* British Psychological Society Developmental Section Annual Conference, Amsterdam.

Danish, J. A., & Saleh, A. (2014). Examining How Activity Shapes Students' Interactions While Creating Representations in Early Elementary Science. *International Journal of Science Education*, *36*(August), 2314–2334. doi:10.1080/09500693.2014.923127

Dar-Nimrod, I., & Heine, S. J. (2011). Genetic essentialism: On the deceptive determinism of DNA. *Psychological Bulletin*, *137*(5), 800–818. doi:10.1037/a0021860

Darwin, C. (1859). *On the Origin of Species*. (W. Bynum, Ed.) (2nd Ed.). London: Penguin Classics.

Dench, G., Ogg, J., & Thomson, K. (1999). *The role of grandparents.* British Social Attitudes: The 16th Report. Published by Ashgate (Aldershot) for the National Centre for Social Research.

Department for Education. (2001). National Curriculum for England: primary science curriculum. London: HMSO.

Department for Education. (2011). Draft National Curriculum 2014: primary science curriculum. London: HMSO.

Department for Education. (2014). National Curriculum for England: primary science curriculum. London: HMSO.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–68. doi:10.1146/annurev-psych-113011-143750

Diamond, A., Barnett, W.S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science, 318*, 1387-1388.

Diamond, A., Kirkham, N., & Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Developmental Psychology*, *38*(3), 352–362. doi:10.1037//0012-1649.38.3.352

Diesendruck, G; Eldror, E. (2011). What Children infer from social categories. *Cognitive Development*, *26*(2), 118–126.

diSessa, A. (1988). Knowledge in Pieces. In G. Forman & P. Pufall (Eds.), *Constructivism in the Computer Age* (pp. 49–70). London: Lawrence Erlbaum Associates, Publishers.

diSessa, A. A. (1983)Phenomenology and evolution of intuition. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 15-33). Hillsdale, NJ: Lawrence Erlbaum Associates.

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction, 10,* 105-225.

diSessa, A. A. (2006). A history of conceptual change research: Threads and fault lines. In R. K. Sawyer (Ed.), the Cambridge handbook of the learning sciences (pp. 265-281). Cambridge: Cambridge University Press.

diSessa, A.A., Gillespie, N., & Esterly, J. (2004). Coherence versus fragmentation in the development of the concept of force. *Cognitive Science, 28*, 843-900.

Dowker, A., & Sigley, G. (2010). Targeted interventions for children with arithmetical difficulties. In R. Cowen, M. Saxton, and A. Tolmie (eds), British Journal of Educational Psychology Monograph Series II: Psychological Aspects of Education – *Current Trends: No. 7. Number Development and Difficulty*. Leicester: BPS.

Driver, R. (1981). Pupils' alternative frameworks in science. *European Journal of Science Education, 3,* 93-101.

Driver, R. (1989). Students' conceptions and the learning of science. *International Journal of Science Education, 11,* 481-490.

Driver, R., Leach, J., Scott, P., & Wood-Robinson, C. (1994). Young People's understanding of science concepts: implications of cross-age studies for curriculum planning. *Studies in Science Education*, *24*(1), 75–100. doi:10.1080/03057269408560040

Driver, R., Guesne, E., & Tiberghien, A. (Eds.). (1985). *Children's ideas in Science*. Milton Keynes: Open University Press.

Duit, R., & Treagust, D. F. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education*, *25*(February), 671–688. doi:10.1080/09500690305016

Duncan, G.L., Brooks-Gunn, J., Klebanov, P. (1994). Economic deprivation and early childhood development. *Child Development, 65*, 296-318.

Dunn, L.M. & Dunn, D.M. (2004). *The British Picture Vocabulary Scale – 3rd Edition (BPVS 3),* London: GL Assessment Limited.

Engel-Clough, E. & Driver, R. (1986). A study in the consistency of the use of students' conceptual frameworks across different task contexts. *Science Education, 70*, 4, 473-496.

Eriksson, H. (1994). Models for knowledge-acquisition tool design. *Knowledge Acquisition*, *6*(1), 47–74. Retrieved from http://www.sciencedirect.com/science/article/B6WMS-45NJHWS-R/2/a9ae9b65fc6bbffb38372c8760466657

Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Senn, T. E. (2004). The contribution of executive functions to emergent mathematic skills in preschool children. *Developmental Neuropsychology*, *26*(1), 465–86. doi:10.1207/s15326942dn2601_6

Evans, E. (2001). Cognitive and contextual factors in the emergence of diverse belief systems: Creation versus evolution. *Cognitive Psychology*, *42*, 217–266.

Fergusson, E., Maughan, B., & Golding, J. (2007). Which children receive grandparent care and what effect does it have? *The Journal of Psychology and Psychiatry, 49(2),* 161-169.

Ferry, A.L., Hespos, S.J., Waxman, S.R. (2013). Nonhuman primate vocalizations support categorization in very young human infants. Proceedings of the National Academy of Sciences of the United States of America, 110(38), 15231-15235.

Fillmore, M.T., & Rush, C.R. (2002). Impaired inhibitory control of behaviour in chronic cocaine users. *Drug and Alcohol Dependency, 66*, 265-273.

Finn, H., Maxwell, M., & Calver, M. (2002). Why does experimentation matter in teaching ecology? *Journal of Biological Education*, *36*(4), 158–162. doi:10.1080/00219266.2002.9655826

Forbus, K., Gentner, D., & Law, K. (1994). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science*, *19*(2), 141–205. doi:10.1016/0364-0213(95)90016-0

Foss, B. New perspectives in child development / edited by Brian Foss. (1974). Harmondsworth : Penguin Education. Retrieved from http://encore.ulrls.lon.ac.uk/iii/encore/record/C__Rb1019252;jsessionid=DC9B61E5CA2379D3DA735B44D28E734C?lang=eng

Fugelsang, J., & Thompson, V. (2003). A dual-process model of belief and evidence interactions in causal reasoning. *Memory and Cognition, 31*, 800-815.

Fuller, T. (2013). Is scientific theory change similar to early cognitive development? Gopnik on science and childhood. *Philosophical Psychology*, *26*(1), 109–128. doi:10.1080/09515089.2011.625114

Galbraith, S. (2011). Accelerated Longitudinal Designs. *Biometrics by the Blowholes*, 4–8.

Ganea, P. a., & Harris, P. L. (2010). Not doing what you are told: Early perseverative errors in updating mental representations via language. *Child Development*, *81*(2), 457–463. doi:10.1111/j.1467-8624.2009.01406.x

Ganea, P. a., Ma, L., & DeLoache, J. S. (2011). Young children's learning and transfer of biological information from picture books to real animals. *Child Development*, *82*(5), 1421–1433. doi:10.1111/j.1467-8624.2011.01612.x

Ganea, P. A., Shutts, K., Spelke, E. S., & DeLoache, J. S. (2007). Thinking of things unseen: Infants' use of language to update mental representations. *Psychological Science, 18,* 734-739.

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in pre-schoolers: a review using an integrative framework. *Psychological Bulletin*, *134*(1), 31–60. doi:10.1037/0033-2909.134.1.31

Gathercole, S. E., Brown, L., & Pickering, S. J. (2003). Working memory assessments at school entry as longitudinal predictors of National Curriculum attainment levels. *Educational and Child Psychology*, *20*(3), 109–122.

Gathercole, S. E., & Pickering, S. J. (2000). Assessment of working memory in six- and seven-year-old children. *Journal of Educational Psychology*, *92*(2), 377–390. doi:10.1037//0022-0663.92.2.377

Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology*, *18*(1), 1–16. doi:10.1002/acp.934

Geary, D.C. (2011). Cognitive predictors of achievement growth in mathematics: a 5 year longitudinal study. *Developmental Psychology*, 47(6), 1539-52.

Gelman, S.A. (2015). From blankies to genes: the role of the non-obvious in children's conceptions of the world. Society for Research in Child Development Biennial Conference, Philadelphia.

Gelman, S. A. (2009). Learning from others: children's construction of concepts. *Annual Review of Psychology*, *60*, 115–40. doi:10.1146/annurev.psych.59.103006.093659

Gelman, S.A. (2003). *The essential child: the origins of essentialism in everyday thought*. New York: Oxford University Press.

Gelman, S.A., & Coley, J.D. (1991). The importance of knowing Dodo is a bird: Categories and inferences in S.A. Gelman & J.P. Byrnes (Eds.), *Perspectives on language and thought: Inter-relations in development* (pp. 146-196). Cambridge: Cambridge University Press.

Gelman, S.A., Coley, J.D., & Gottfried, G.M. (1994). Essentialist beliefs in children: the acquisition of concepts and theories. In L.A. Hirschfeld & S.A. Gelman (Eds.). *Mapping the Mind: domain specificity in cognition and culture,* (chapter 13). Cambridge: Cambridge University Press.

Gelman, S. A., & Hirschfeld, L. A. (1999). How biological is essentialism? In S. Altran & D. Medin (Eds.), *Folk Biology* (chapter 12). Cambridge, Massachusetts, USA: MIT.

Gelman, S.A., & Kremer, K.E. (1991). Understanding natural cause: children's explanations of how objects and their properties originate. *Child Development, 62(2*), 396-414.

Gelman, S.A., & Markman, E.A. (1986). Categories and induction in young children. *Cognition, 23(3),* 183-209.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, *52*(1), 45–56. doi:10.1037//0003-066X.52.1.45

Georghiades, P. (1999). *Conceptual change learning in primary science: A step forward? - British Education Index - ProQuest Dialog*. *Monograph, conference paper*. Retrieved from http://search.proquest.com/professional/briteducationindex/docview/771758334/13A21E3EDA64BA34587/18?accountid=27115

Ghazali, Z. (2015). Can we be scientific about science education? *The Psychologist, 28*, 992-993.

Ghazali, Z., & Tolmie, A. (2014). New approaches to understanding the development of biological concepts in young children. *Educacio Siglo XX1, 32*(2), 97-118.

Glass, R., & Oliveira, A. W. (2014). Science Language Accommodation in Elementary School Read-Alouds. *International Journal of Science Education*, *0*(August), 1–33. doi:10.1080/09500693.2013.802057

Goldberg, R. F., & Thompson-Schill, S. L. (2009). Developmental “Roots” in Mature Biological Knowledge. *Psychological Science*, *20*(4), 480–487. doi:10.1111/j.1467-9280.2009.02320.x

Gopnik, A. (1996). The scientist as child. *Philosophy of Science, 63,* 485-514.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A Theory of Causal Learning in Children: Causal Maps and Bayes Nets. *Psychological Review*, *111*(1), 3–32. doi:10.1037/0033-295X.111.1.3

Gopnik, A., Meltzoff, A. N., & Kuhl, P. K. (2001). The scientist in the crib: What early learning tells us about the mind. New York, NY: Harper Collins.

Gopnik, A., & Schulz, L. (2004). Mechanisms of theory formation in young children. *Trends in Cognitive Sciences*, *8*(8), 371–7. doi:10.1016/j.tics.2004.06.005

Gopnik, A., & Welman, H. (1994). The Theory Theory. In L.A. Hirschfeld and S.A. Gelman (Eds.). *Mapping the Mind: domain specificity in cognition and culture* (p.257). Cambridge: Cambridge University Press.

Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: Two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology*, *37*(5), 620–629. doi:10.1037//0012-1649.37.5.620

Grant, D.S., & Berg, E.A. (1948). A behavioural analysis of degree of reinforcement and ease of shifting to new responses in a weight-type card sorting problem. *Journal of Experimental Psychology, 321,* 404-411.

Gregory, E., Long, S., & Volk, D. (2004). Many pathways to literacy: young children learning with siblings, grandparents, peers and communities. London: Routledge.

Grotzer, T. A., & Basca, B. B. (2003). How does grasping the underlying causal structures of ecosystems impact students’ understanding? *Journal of Biological Education*, *38*(1), 16–29. doi:10.1080/00219266.2003.9655891

Grotzer, T. a. (2003). Learning to Understand the Forms of Causality Implicit in Scientifically Accepted Explanations. *Studies in Science Education*, *39*(1), 1–74. doi:10.1080/03057260308560195

Gülgöz, S., & Gelman, S. a. (2014). Children’s recall of generic and specific labels regarding animals and people. *Cognitive Development*, *33*, 84–98. doi:10.1016/j.cogdev.2014.05.002

Ha, M., & Nehm, R. H. (2013). Darwin’s Difficulties and Students' Struggles with Trait Loss: Cognitive-Historical Parallelisms in Evolutionary Explanation. *Science and Education*, 1–24. doi:10.1007/s11191-013-9626-1

Hair, J.F., Black, W.C., Babin, B.J., & Anderson, R.E. (2009). Multivariate data analysis. Upper Saddle River, NJ: Prentice Hall.

Harlen, W. (2001). Research in primary science education. *Journal of Biological Education*, *35*(2), 61–65. doi:10.1080/00219266.2000.9655743

Harold, G. T., Leve, L. D., Kim, H. K., Mahedy, L., & Gaysina, D. (2014). Maternal caregiving and girls' depressive symptoms and antisocial behaviour trajectories: An examination among high-risk youth. *Developmental Psychopathology*, *26*(0), 1461–1475. doi:10.1017/S095457941400114X.Maternal

Harris, P. (1994). Thinking by children and scientists: False analogies and neglected similarities. In. L.A. Hirschfeld, & S.A. Gelman (Eds.). *Mapping the Mind: domain specificity in cognition and culture,* (chapter 11). Cambridge: Cambridge University Press.

Harris, P. L. (2002). What do children learn from testimony? In P. Carruthers, S. Stich, & M. Siegal (Eds.), *The cognitive basis of science* (pp. 316-334). Cambridge: Cambridge University Press.

Harris, P.L., & Koenig, M.A. (2006). Trust in Testimony: how children learn about science and religion. *Child Development, 77(3*), 505-524.

Hart, B., & Risley, T.R. (1995). *Meaningful Differences in the Everyday Experience of Young American Children.* Baltimore, MD; Brookes.

Hatano, G. (1990). The nature of everyday science: A brief introduction. *British Journal of Developmental Psychology, 8*, 245-250.

Hatano, G., & Inagaki, K. (1994). Young children's naive theory of biology. *Cognition*, *50*(1-3), 171–188. Retrieved from http://www.sciencedirect.com/science/article/B6T24-45WHV4K-3P/2/677bd0c39c417f5096fe1bfcb1d1e753

Hecht, S.A. (2002). Counting on working memory in simple arithmetic when counting is used for problem solving. *Memory and Cognition, 30(3),* 447-55.

Hecht, S.A., Torgesen, J.K., Wagner, R.K., & Rashotte, C.A. (2001). The relations between phonological processing abilities and emerging individual differences in mathematical computation skills: a longitudinal study from second to fifth grades. *Journal of Experimental Child Psychology, 79(2),* 192-227.

Herrmann, P. A., French, J. A., DeHart, G. B., & Rosengren, K. S. (2013). Essentialist Reasoning and Knowledge Effects on Biological Reasoning in Young Children. *Merrill-Palmer Quarterly*, *59*(2), 198–220. doi:10.1353/mpq.2013.0008

Hipkins, R., Bull, A., & Joyce, C. (2008). The interplay of context and concepts in primary school children's systems thinking. *Journal of Biological Education*, *42*(2), 73–77. doi:10.1080/00219266.2008.9656114

Hirschfeld, L.A., & Gelman, S.A. (1994). *Mapping the Mind: domain specificity in cognition and culture* (Eds.). Cambridge: Cambridge University Press.

Hoff-Ginsberg, E. (1991). Mother-child conversation in different social classes and communicative settings. Child Development, 62, 782-786.

Howe, C. (1998). *Conceptual Structure in Childhood and Adolescence*. London: Routledge.

Howe, C., McWilliam, D., & Cross, G. (2005). Chance favours only the prepared mind: Incubation and the delayed effects of peer collaboration. *British Journal of Psychology (London, England : 1953)*, *96*(Pt 1), 67–93. doi:10.1348/000712604X15527

Howe, C., Nunes, T., & Bryant, P. (2010). Intensive quantities: Why they matter to developmental research. *British Journal of Developmental Psychology*, *28*, 307–329.

Howe, C., Taylor Tavares, J., Devine, A. (2010). Children's conceptions about physical events: Explicit and tacit understanding of horizontal motion. Manuscript in preparation.

Howe, C., Tavares, J., & Devine, A. (2012). Everyday conceptions of object fall: Explicit and tacit understanding during middle childhood. *Journal of Experimental Child Psychology*, *111*, 351–366.

Howe, C., Tolmie, A., Rodgers, C. (1992). The acquisition of conceptual knowledge in science by primary school children: Group interaction and the understanding of motion down an incline. *British Journal of Developmental Psychology, 10,* 113-130.

Howe, C., Tolmie, A., & Sofroniou, N. (1999). Experimental appraisal of personal beliefs in science: constraints on performance in the 9 to 14 age group. British Journal of Educational Psychology, 69, 243-274.

Howe, C., Tolmie, A., Thurston, A., Topping, K., Christie, D., Livingston, K., Jessiman, E., & Donaldson, C. (2007). Group work in elementary science: towards organisational principles for supporting pupil learning. Learning and Instruction, 17, 549-563.

Huizinga, M., Dolan, C. V, & van der Molen, M. W. (2006). Age-related change in executive function: developmental trends and a latent variable analysis. *Neuropsychologia*, *44*(11), 2017–36. doi:10.1016/j.neuropsychologia.2006.01.010

Hulme, C., & Snowling, M.J. (2009). *Developmental Disorders of Language Learning and Cognition*. Chichester: Wiley-Blackwell.

Hyde, J. S., & Linn, M. C. (2006). Gender Similarities in Mathematics and Science Evidence for Gender Similarities. *October*, *314*(October), 599–600.

Inagaki, K. (1997). Emerging distinctions between naive biology and naive psychology. In H. Wellmann & K. Inagaki (Eds.), *The emergence of core domains of thought Childrens reasoning about physical psychological and biological phenomena* (pp. 27–44). Jossey-Bass Publishers.

Inagaki, K., & Hatano, G. (2008). Conceptual change in naive biology. In S. Vosniadou (Ed.), *International Handbook of Research on Conceptual Change* (pp. 240–262). Routledge.

Inagaki, K., & Hatano, G. (2004). Vitalistic causality in young children's naive biology. *Trends in Cognitive Sciences*, *8*(8), 356–362. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15335462

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence* (A. Parsons & S. Milgram, Trans.). London: Routledge & Kegan Paul.

Janssen, F., & Waarlo, A. J. (2010). Learning Biology by Designing. *Journal of Biological Education*, *44*(2), 88–92. doi:10.1080/00219266.2010.9656199

Johnson, S. C., & Carey, S. (1998). Knowledge enrichment and conceptual change in folkbiology: evidence from Williams syndrome. *Cognitive Psychology*, *37*(2), 156–200. doi:10.1006/cogp.1998.0695

Jordan, R., & Duncan, R. G. (2009). Student teachers' images of science in ecology and genetics. *Journal of Biological Education*, *43*(2), 62–69. doi:10.1080/00219266.2009.9656153

Kalish, C. (1996). Causes and Symptoms in Preschoolers' Conceptions of Illness. *Child Development*, *67*(4), 1647–1670. doi:10.1111/j.1467-8624.1996.tb01819.x

Kalish, C. W., Rogers, T. T., Lang, J., & Zhu, X. (2011). Can semi-supervised learning explain incorrect beliefs about categories? *Cognition*, *120*(1), 106–118. doi:10.1016/j.cognition.2011.03.002

Kallai, A. K., & Reiner, M. (2010). The Source of Misconceptions in Physics: When Event-Related Potential Components N400 and P600 Disagree. Pittsburgh.

Karmiloff-Smith, A. (1988). The child is a theoretician not an inductivist. *Mind & Language, 3,* 183-195.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences*, *2*(10), 389–398. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/21227254

Karmiloff-Smith, A. (1992). *Beyond Modularity* (2nd Ed.). Cambridge, Massachusetts, USA: Massachusetts Institute of Technology Press.

Karmiloff-Smith, A., & Inhelder, B. (1974). "If you want to get ahead, get a theory." *Cognition, 3*, 195-212.

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Jiménez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition*, *116*(3), 321–340. doi:10.1016/j.cognition.2010.05.011

Kawalkar, A., & Vijapurkar, J. (2011). Scaffolding Science Talk: The role of teachers' questions in the inquiry classroom. *International Journal of Science Education*, (June 2015), 1–24. doi:10.1080/09500693.2011.604684

Keil, F.C. (1996). *Concepts, kinds, and cognitive development.* Cambridge, MA: MIT Press.

Keil, F. C. (1987). Conceptual development and category structure. In U. Neisser (Ed.), *Concepts and conceptual development Ecological and intellectual factors in categorization Emory symposia in cognition 1* (pp. 175–200). Cambridge University Press.

Keil, F.C. (1994). The birth and nurturance of concepts by domain: the origins of concepts of living things. In. L.A. Hirschfeld, & S.A. Gelman (Eds.). *Mapping the Mind: domain specificity in cognition and culture,* (p.234). Cambridge: Cambridge University Press.

Keil, F. C., & Lockhart, K. L. (1999). Explanatory understanding in conceptual development. In E. K. Scholnick & K. Nelson (Eds.), *Conceptual development Piagets legacy The Jean Piaget Symposium series* (pp. 103–130 ST – Explanatory understanding in concept). Lawrence Erlbaum Associates, Publishers.

Keleman, D. (2004). Are children “intuitive theists”? Reasoning about purpose and design in nature. *Psychological Science, 15,* 295-301.

Kelemen, D., Emmons, N. a., Seston Schillaci, R., & Ganea, P. A. (2014). Young Children Can Be Taught Basic Natural Selection Using a Picture-Storybook Intervention. *Psychological Science*, *25*(February), 893–902. doi:10.1177/0956797613516009

Keleman, D., Rottman, J., Seston, R. (2013). Professional physical scientists display tenacious teleological tendencies: purpose-based reasoning as a cognitive default. *Journal of Experimental Psychology, 142(4),* 1074-1083.

Kemp, C., Shafto, P., & Tenenbaum, J. B. (2012). An integrated account of generalization across objects and features. *Cognitive Psychology*, *64*(1-2), 35–73. doi:10.1016/j.cogpsych.2011.10.001

Kenner, C., Ruby, M., Jessel, J., Gregory. E., & Arju, T. (2007). Intergenerational learning between children and grandparents in East London. Journal of Early Childhood Research, 5(3), 219-243.

Keynes, R. (2009). Darwin’s ways of working — the opportunity for education. *Journal of Biological Education*, *43*(3), 101–103. doi:10.1080/00219266.2009.9656162

Kinchin, I. M. (2003). Effective teacher ↔ student dialogue: a model from biological education. *Journal of Biological Education*, *37*(3), 110–113. doi:10.1080/00219266.2003.9655864

Kirkham, N. Z., Slemmer, J.A., & Johnson, S.P. (2002). Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition, 83,* B35-B42.

Kirkpatrick, L. A. (1999). Toward an Evolutionary Psychology of Religion and Personality. *Journal of Personality*, *67*(6), 921–952. doi:10.1111/1467-6494.00078

Kloos, H., & Sloutsky, V. M. (2008). What’s behind different kinds of kinds: effects of statistical density on learning and representation of categories. *Journal of Experimental Psychology. General*, *137*(1), 52–72. doi:10.1037/0096-3445.137.1.52

Koerber, S., Sodian, B., Thoermer, C., Ulrike, N. (2005). Scientific reasoning in young children: pre-schoolers’ ability to evaluate covariation evidence. *Swiss Journal of Psychology, 64(3*), 141-152.

Korfiatis, K. J., & Tunnicliffe, S. D. (2012). The living world in the curriculum: ecology, an essential part of biology learning. *Journal of Biological Education*, *46*(3), 125–127. doi:10.1080/00219266.2012.715425

Krippendorff, K. (1980). *Content analysis: An Introduction to its methodology.* London: Sage Publications Ltd.

Krombaβ, A., & Harms, U. (2008). Acquiring knowledge about biodiversity in a museum — are worksheets effective? *Journal of Biological Education*, *42*(4), 157–163. doi:10.1080/00219266.2008.9656134

Kuhn, T.S. (1962). *The structure of scientific revolutions*, Chicago, IL: University of Chicago Press.

Kuhn, D., Amsel, E., & O'Loughin, M. (1988). The development of scientific thinking skills. Florida: Academic Press Orlando.

Kuperman, V., Stradthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. Behavioural Research Methods, 44(4), 978-990.

Kushnir, T., Xu, F., & Wellman, H.M. (2010). Young children use statistical sampling to infer the preferences of others. *Psychological Science, 21,* 1134-1140.

Kwon, Y., & Lawson, A.E. (2000). Linking brain growth with the development of scientific reasoning ability and conceptual change during adolescence. Journal of Research in Science Teaching, 37, 44-62.

Lappi, O. (2012). Qualitative Quantitative and Experimental Concept Possession, Criteria for Identifying Conceptual Change in Science Education. *Science & Education*, *22*(6), 1347–1359. doi:10.1007/s11191-012-9459-3

Lawson, A. E. (1988). The acquisition of biological knowledge during childhood: Cognitive conflict or tabula rasa? *Journal of Research in Science Teaching*, *25*(3), 185–199. doi:10.1002/tea.3660250304

Leach, J., Driver, R., Scott, P., & Wood-Robinson, C. (1995). Children's ideas about ecology 1: theoretical background, design and methodology. *International Journal of Science Education*, *17*(6), 721–732. doi:10.1080/0950069950170604

Leach, J., Driver, R., Scott, P., & Wood-Robinson, C. (1996). Children's ideas about ecology 2: ideas found in children aged 5-16 about the cycling of matter. *International Journal of Science Education*, *18*(1), 19–34. doi:10.1080/0950069960180102

Legare, C. H., Wellman, H. M., & Gelman, S. a. (2009). Evidence for an explanation advantage in naïve biological reasoning. *Cognitive Psychology*, *58*(2), 177–94. doi:10.1016/j.cogpsych.2008.06.002

Lehto, J., Juujarvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*, 59–80.

Lewis, J. (2009). Can theoretical constructs in science be generalised across disciplines? *Journal of Biological Education*, *44*(1), 5–11. doi:10.1080/00219266.2009.9656185

Lewis, J., Driver, R., Leach, J., & Wood-Robinson, C. (1997). *Young people's understanding of and attitudes towards to the 'new genetics' project.* Working paper 2: understanding of basic genetics and DNA technology. Centre for Science and Mathematics Education, Learning in Science Research Group, Leeds: University of Leeds.

Linquist, S., Machery, E., Griffiths, P. E., & Stotz, K. (2011). Exploring the folkbiological conception of human nature. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1563), 444–53. doi:10.1098/rstb.2010.0224

Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology*, *50*(4), 315–53. doi:10.1016/j.cogpsych.2004.09.004

Luciana, M., & Nelson, C. (1998). The functional emergence of prefrontally-guided working memory systems in four- to eight-year-old children. *Neuropsychologia*, *36*(3), 273–293.

Mareschal, D., & Quinn, Paul, C. (2001). Categorization infancy. *Trends in Cognitive Sciences*, *5*(10), 443–450.

Mareschal, C., Johnson, M. H., & Grayson, A. (2004). Brain and cognitive development. In J. Oates & Grayson (Eds.), *Cognitive and language development in children* (pp.113-161). Oxford: Blackwell Publishing.

Mason, R., & Just, M. (2015). Physics instruction induces changes in neural knowledge representation during successive stages of learning. *NeuroImage*, (111), 36–48.

Maughan, B., Pickles, A., Rowe, R., Costello, E. J., & Angold, A. (2000). Developmental Trajectories of Aggressive and Non-Aggressive Conduct Problems. *Journal of Quantitative Criminology*, *16*(2), 199–221. doi:10.1023/a:1007516622688

Maurice-Naville, D., & Montangero, J. (1992). The development of diachronic thinking: 8-12-year-old children's understanding of the evolution of forest disease. *British Journal of Developmental Psychology*, *10*(4), 365–383. doi:10.1111/j.2044-835X.1992.tb00583.x

Mayer, R. E. (2002). Understanding conceptual change: A commentary. In M. Limon & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice* (pp.101-114). Dordrect: Kluwer Academic.

Mayer, D., Sodian, B., Koerber, S., & Schwippert, K. (2014). Scientific reasoning in elementary school children: Assessment and relations with cognitive abilities. *Learning and Instruction*, *29*, 43–55. doi:10.1016/j.learninstruc.2013.07.005

McCloskey, M., & Kohl, D. (1983). Naïve physics: The curvilinear impetus principle and its role in interaction with moving objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 146-156.

Melhuish, E., Sylva., K., Sammons, P., Siraj-Blatchford, I. & Taggart, B. (2001). *The Effective Provision of Pre-school Education Project, Technical Paper 7: Social/behavioural and Cognitive Developmental at 3-4 years in relation to family background. London:* Institute of Education/DfEE.

Mercy, J.A., & Steelman, L.C. (1982). Familial influence on the intellectual attainment of children. *American Sociological Review, 47*, 532-542.

Miller Singley, A. T., & Bunge, S. a. (2014). Neurodevelopment of relational reasoning: Implications for mathematical pedagogy. *Trends in Neuroscience and Education*, *3*(2), 33–37. doi:10.1016/j.tine.2014.03.001

Miyake, A., Friedman, N. P., Emerson, M.J., Witzki, A.J., & Howerter, A. (2000). *The unity and diversity of executive functions and their contributions to complex frontal lobe tasks: A latent variable analysis.* Cognitive Psychology, 41, 49-100.

Montangero, Jacques; Maurice-Naville, D. (1997). *Piaget or the advance of knowledge*. New Jersey; London: Lawrence Erlbaum Associates, Publishers.

Montangero, J., Pons, F., & Cattin, J.-P. (2000). The diachronic approach and solutions to interpersonal conflicts. *British Journal of Developmental Psychology*, *18*(3), 415–429. doi: 10.1348/026151000165779

Myant, K.A., & Williams, J.M. (2005). Children's concepts of health and illness: Understanding of contagious illnesses, non-contagious illnesses and injuries. *Journal of Health Psychology, 10(6*), 805-819.

Nayfeld, I., Fuccillo, J., & Greenfield, D. B. (2013). Executive functions in early learning: Extending the relationship between executive functions and school readiness to science. *Learning and Individual Differences*, *26*, 81–88. doi:10.1016/j.lindif.2013.04.011

Nersessian, N.J. (2003). Kuhn, conceptual change, and cognitive science. In T. Nickles (Ed.), Thomas Kuhn (pp.178-211). Cambridge: Cambridge University Press.

Nunes, T., Bryant, P., Barros, R., Sylva, K. (2012). The relative importance of two different mathematical abilities to mathematical achievement. British Journal of Educational Psychology. 82(1), 136-56.

Okamato, Y., & Case, R. (1996). Exploring the microstructure of children's central conceptual structures in the domain of number. *Monographs of the society for research in child development,* serial No. 246, 61(1-2).

Okur, E. (2011). The common methods used in biodiversity education by primary school teachers, *7*(1), 142–159.

Ojalehto, B., Waxman, S.R., & Medin, D.L. (2013). Teleological reasoning about nature: intentional design or relational perspectives? Trends in Cognitive Sciences, 17(4), 166-171.

Opfer, J. E., Nehm, R. H., & Ha, M. (2012). Cognitive foundations for science assessment design: Knowing what students know about evolution. *Journal of Research in Science Teaching*, *49*(6), 744–777. doi:10.1002/tea.21028

Ozdemir, G., & Clark, D.B. (2007). An overview of conceptual change theories. *Eurasia Journal of Mathematics, Science & Technology Education*, *3*(4), 351–361.

Ozdemir, G., & Clarke, D.B. (2009). Coherence versus fragmentation: A study of Turkish students' understanding of force. *Journal of Research in Science Teaching, 46(5),* 570-596.

Parcel, T.L., & Menaghan, E.G. (1990). Maternal working conditions and children's verbal facility: studying the intergenerational transmission of inequality from mothers to young children. *Social Psychology Quarterly, 53,* 132-147.

Phillips, S. & Tolmie, A. (2007). Children's performance on and understanding of the balance scale problem: the effects of parental support. *Infant and Child Development, 16,* 95-117.

Piaget, J. (1971). *Biology and Knowledge*. Edinburgh: Edinburgh University Press.

Piaget, J. (1972). *Jean Piaget's Psychology and Epistemology: Towards a Theory of Knowledge.* (P. A. Wells, Ed.). Suffolk, UK: Penguin Univeristy Books.

Piaget, J. (1985). *The Equilibration of Cognitive Structures*. Chicago: Chicago University Press.

Pickering, S., & Gathercole, S.E. (2001). Working memory test battery for children (WMTB-C). London: Psychological Corporation.

Poirel, N., Borst, G., Simon, G., Rossi, S., Cassotti, M., Pineau, A., & Houdé, O. (2012). Number conservation is related to children's prefrontal inhibitory control: an fMRI study of a piagetian task. *PloS One*, *7*(7), e40802. doi:10.1371/journal.pone.0040802

Pons, F., & Montangero, J. (1999). Is diachronic thought a specific reasoning ability? *Swiss Journal of Psychology*, *58*(3), 191–200. doi:10.1024//1421-0185.58.3.191

Posner, George. J; Strike, Kenneth. A; Hewson, Peter. W; Gertzog, W. A. (1982). Accommodation of a Scientific Conception : Toward a Theory of Conceptual Change *. *Science Education*, *66*(1968), 211–227.

Prince-William, S., Douglass, P. T., Gordon, J., William, R., Ramirez, E., & Manuel, J.S. (1969). Skill and Conservation: A study of pottery-making children. *Developmental Psychology, 16* (6, part 1).

Quine, W.V.O. (1960). *Word and Object*. Cambridge, MA: MIT Press.

Ramsden, S., Richardson, F., Josse, G., Thomas, M., Ellis, C., Shakeshaft, C., Mohamed, L.S., & Price, C.J. (2011). Verbal and non-verbal intelligence changes in the teenage brain. *Nature*, *479*, 113-116. (3 November 2011) doi:10.1038/nature10514

Rashid, K., Sanaulla, R., Iqbal, M.Z., & Khalid, N. (2013). Pre-school attendees and non-preschool attendees' academic achievement and social skills. Interdisciplinary Journal of Contemporary Research in Business, 4(9), 1146-1157.

Rattermann, M. J., & Gentner, D. (1998). More evidence for a relational shift in the development of analogy: Children's performance on a causal-mapping task. *Cognitive Development*, *13*, 453–478.

Rhodes, M., & Gelman, S. A. (2009). A developmental examination of the conceptual structure of animal, artifact, and human social categories across two cultural contexts. *Cognitive Psychology*, *59*(3), 244–74. doi:10.1016/j.cogpsych.2009.05.001

Rhodes, M., & Wellman, H. (2013). Constructing a new theory from old ideas and new evidence. *Cognitive Science, 37(3*), 592-604.

Rigney, J. C., & Callanan, M. A. (2011). Patterns in parent–child conversations about animals at a marine science center. *Cognitive Development*, *26*(2), 155–171. doi:10.1016/j.cogdev.2010.12.002

Romine, C.B., Lee, D., Wolfe, M.E., Homacj, S., George, C., & Riccio, C.A. (2004). Wisconsin card sorting test with children: a meta-analytic study of sensitivity and specificity. Archives of Clinical Neuropsychology, 19(8), 1027-1041.

Rose, S. (2001). Revisiting evolutionary psychology and psychiatry. *The British Journal of Psychiatry*, *179*(6), 558–a–558. doi:10.1192/bjp.179.6.558-a

Rose, S. (2009). Darwin, race and gender. *EMBO Reports*, *10*(4), 297–8. doi:10.1038/embor.2009.40

Rowlands, M. (2001). The development of children's biological understanding. *Journal of Biological Education*, *35*(2), 66–68. doi:10.1080/00219266.2000.9655744

Royal Society. (2010). *Science and Mathematics Education 5-14: A state of the nation report*. London: Royal Society.

Royal Society. (2011). *Neuroscience: Implications for education and lifelong learning*. London: Royal Society.

Sander, E., Jelemenská, P., & Kattmann, U. (2006). Towards a better understanding of ecology. *Journal of Biological Education*, *40*(3), 119–123. doi:10.1080/00219266.2006.9656028

Scarr, S., & Weinberg, R.A. (1978). The influence of family background on intellectual attainment. *American Sociological Review, 43,* 674-692.

Schilders, M., Sloep, P., Peled, E., & Boersma, K. (2009). Worldviews and evolution in the biology classroom. *Journal of Biological Education*, *43*(3), 115–120. doi:10.1080/00219266.2009.9656165

Schulz, L. E., & Gopnik, A. (2004). Causal learning across domains. *Developmental Psychology*, *40*(2), 162–76. doi:10.1037/0012-1649.40.2.162

Schulz, L. E., Gopnik, A., & Glymour, C. (2007). Preschool children learn about causal structure from conditional interventions. *Developmental Science*, *10*(3), 322–32. doi:10.1111/j.1467-7687.2007.00587.x

Sellers, R., Harold, G. T., Elam, K., Rhoades, K. a., Potter, R., Mars, B., & Collishaw, S. (2014). Maternal depression and co-occurring antisocial behaviour: Testing maternal hostility and warmth as mediators of risk for offspring psychopathology. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *55*, 112–120. doi:10.1111/jcpp.12111

Shtulman, A. (2006). Qualitative differences between naïve and scientific theories of evolution. *Cognitive Psychology*, *52*(2), 170–94. doi:10.1016/j.cogpsych.2005.10.001

Shtulman, A., & Schulz, L. (2008). The relation between essentialist beliefs and evolutionary reasoning. *Cognitive Science*, *32*(6), 1049–1062.

Shtulman, A., & Valcarcel, J. (2012). Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition*, *124*, 209–215. doi:cognition.2012.04.005

Siegler, R. S. (2000). The Rebirth of Children's Learning. *Child Development*, *71*(1), 26–35. doi:10.1111/1467-8624.00115

Skipper, J.I. (2015). *The NOLB model: a model of the natural organization of language and the brain*. Invited address, Centre for Educational Neuroscience seminar series, London, UK.

Slingsby, D. (2009). Charles Darwin, Biological Education and diversity: past present and future. *Journal of Biological Education*, *43*(3), 99–100. doi:10.1080/00219266.2009.9656161

Slaughter, V., & Gopnik, A. (1996). Conceptual coherence in the child's theory of mind: Training children to understand belief. *Child Development, 67*, 6, 2567-2988.

Slaughter, V., Jaakkola, R., & Carey, S. (1999). Constructing a coherent theory: Children's biological understanding of life and death. In M. Siegal & C. Peterson (Eds.). *Children's understanding of biology and health* (pp.71-96). Cambridge: Cambridge Univeristy Press.

Smith, K. V., Loughran, J., Berry, A., & Dimitrakopoulos, C. (2012). Developing Scientific Literacy in a Primary School. *International Journal of Science Education*, *34*(1), 127–152. doi:10.1080/09500693.2011.565088

Snaddon, J. L., Turner, E. C., & Foster, W. A. (2008). Children's perceptions of rainforest biodiversity: which animals have the lion's share of environmental awareness? *PloS One*, *3*(7), 5. doi:10.1371/journal.pone.0002579

Solomon, G. E. A. (2002). Birth, kind and naïve biology. *Developmental Science*, *5*(2), 213–218. doi:10.1111/1467-7687.00223

Soltesz, F., Szucs, D., & Szucs, L. (2010). Relationships between magnitude representation, counting and memory in 4- to 7-year-old children: a developmental study. Behavioural and Brain Functions, 6:13. Online www.behavioralandbrainfunctions.com/content/6/1/13 (accessed 11th October 2014).

Southerland, S.A., Abrams, E., Cummins, C.L., & Anzelmo, J. (2001). Understanding students' explanations of biological phenomena: conceptual frameworks of pprims. *Science Education, 85,* 311-327.

Spelke, E. S. (1994). Initial knowledge: Six suggestions. *Cognition, 50*, 431-445.

Spelke, E. S., & Kinzler, K. D. (2000). Core knowledge. *Developmental Science*, *10*(1), 1233–1243. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17181705

Springer, K. (1996). Children's understanding of a biological basis for parent-offspring relations. *Child Development*, *67*(6), 2841–2856.

Springer, K. (1999). Acquiring a Theory of Biology. In M. Siegal & C. Peterson (Eds.), *Children's Understanding of Biology and Health* (pp. 45–70). Cambridge, UK: Cambridge University Press.

Standards and Testing Agency. (2013). *National Curriculum assessments: past papers.* Retrieved 11th June, 2015 from https://www.gov.uk/government/collections/key-stage-2-tests-past-papers

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology (2006)*, *59*(4), 745–59. doi:10.1080/17470210500162854

Stoet, G., & Geary, D. C. (2013). Sex Differences in Mathematics and Reading Achievement Are Inversely Related: Within- and Across-Nation Assessment of 10 Years of PISA Data. *PLoS ONE*, *8*(3), e57988. doi:10.1371/journal.pone.0057988

Strevens, M. (2000). The essentialist aspect of naive theories. *Cognition*, *74*, 149–175.

Szatmari, P., Georgiades, S., Duku, E., Bennett, T. a., Bryson, S., Fombonne, E., & Thompson, A. (2015). Developmental Trajectories of Symptom Severity and Adaptive Functioning in an Inception Cohort of Preschool Children With Autism Spectrum Disorder. *JAMA Psychiatry*. doi:10.1001/jamapsychiatry.2014.2463

Tarlowski, A. (2006). If it’s an animal it has axons: Experience and culture in preschool children's reasoning about animates. *Cognitive Development*, *21*(3), 249–265. doi:10.1016/j.cogdev.2006.02.001

Tolmie, A. (2011). *Language and causal reasoning in science*. Invited address, Brain, Neurosciences, and Education SIG, American Educational Research Association Annual Meeting, New Orleans

Tolmie, A. (2012).*Understanding core skills and influences in primary school science learning: Taking a scientific approach.* London: IOE Press.

Tolmie, A., Tenenbaum, H, & Pino-Pasternak, D. (2009). Generalisation in children’s science explanations. European Association for Learning and Instruction Biennial Conference, Amsterdam.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behavioral and Brain Sciences*, *28*(5), 675–691; discussion 691–735. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16262930

Turkheimer, E. (2011). Genetics and human agency: Comment on Dar-Nimrod and Heine (2011). *Psychological Bulletin*, *137*(5), 825–828. doi:10.1037/a0024306

Vosniadou, S. (2002). Mental models in conceptual change research. In L. Magnani & N. J. Nersessian (Eds.). Model-based reasoning: science, technology, values (pp.353-368). New York: Kluwer Academic/Plenum.

Vosniadou, S. (2014). Examining cognitive development from a conceptual change point of view: The framework theory approach. *European Journal of Developmental Psychology*, (June 2014), 1–17. doi:10.1080/17405629.2014.921153

Vosnoadou, S., & Brewer, W.F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology, 24*, 535-585.

Vosniadou, S., & Ioannides, C. (1998). From conceptual development to science education: a psychological point of view. *International Journal of Science Education*, *20*(10), 1213–1230. doi:10.1080/0950069980201004

Vygotsky, L.S. (1962). *Thought and Language.* Cambridge, MA: MIT press.

Vygotsky, L.S. (1978). *Mind in Society: the development of higher psychological processes.* Cambridge: Harvard University Press.

Wagoner, B., & Jensen, E. (2010). Science Learning at the Zoo : Evaluating Children ' s Developing Understanding of Animals and their Habitats. *Psychology and Society*, *3*(1), 65–76.

Wasmann-Frahm, A. (2009). Conceptual change through changing the process of comparison. *Journal of Biological Education*, *43*(2), 71–77. Retrieved from https://login.elibrary.ioe.ac.uk/login?url=http://search.proquest.com/professional/docview/772137508?accountid=27115

Waxman, S., Medin, D., & Ross, N. (2007). Folkbiological reasoning from a cross-cultural developmental perspective: early essentialist notions are shaped by cultural beliefs. *Developmental Psychology*, *43*(2), 294–308. doi:10.1037/0012-1649.43.2.294

Wechsler, D. (1999).Manual for the Wechsler abbreviated intelligence scale (WASI). San Antonio, Texas: The Psychological Corporation.

Wellcome Trust. (2013). *Perspectives on Education: Effects from accountabilities*. Retrieved from http://www.wellcome.ac.uk/About-us/Publications/Reports/Education/Perspectives/stellent/groups/corporatesite/@msh_peda/documents/web_document/WTP052346.pdf

Wellcome Trust. (2005). *Primary Horizons: Starting out in Science*. London.

Wellman, H.M., & Gelman, S.A. (1998). Knowledge acquisition in foundational domains. In D. William (Ed.), Handbook of child psychology: Volume 2: Cognition, perception, and language, (pp.523-573). Hoboken, NJ, US: John Wiley & Sons Inc, xxvi, 1030pp.

West, L., & Pines, L. (1984). An interpretation of research in conceptual understanding within a sources-of-knowledge framework. *Research in Science Education, 14*, 47-56.

West, L., & Pines, L. (1986). Conceptual understanding and science learning: An interpretation of research within a sources-of-knowledge framework. *Science Education, 70,* 583-604.

Williams, J.M., & Binnie, L.M. (2002). Children;s concepts of illness: An interviention to improve knowledge. British Journal of Health Psychology, 7, 129-147.

Williams, J. M., & Smith, L. A. (2006). Social and experiential influences on the development of inheritance concepts. *International Journal of Behavioral Development*, *30*(2), 148–157. doi:10.1177/0165025406063630

Williams, J. M. (2012). Children and adolescents' understandings of family resemblance: a study of naïve inheritance concepts. *The British Journal of Developmental Psychology*, *30*(Pt 2), 225–52. doi:10.1111/j.2044-835X.2011.02031.x

Williams, J. M., & Tolmie, A. K. (2000). Conceptual change in biology: Group interaction and the understanding of inheritance. *British Journal of Developmental Psychology*, *10*(4), 625–649.

Williams, J. M., & Affleck, G. (1999). The Effects of an Age-appropriate Intervention on Young Children's Understanding of Inheritance. *Educational Psychology*, *19*(3), 259–275. doi:10.1080/0144341990190302

Williams, J. M., & Smith, L. A. (2010). Concepts of kinship relations and inheritance in childhood and adolescence. *British Journal of Developmental Psychology*, *28*(3), 523–546. doi:10.1348/026151009X449568

Willoughby, M.T., Blair, C.B., Wirth RJ., & Greenberg, M. (2012). The measurement of executive function at age 5: psychometric properties and relationship to academic achievement. Psychological Assessment, 24 (1), 226-239.

Wilson, R. (2009). The demand for STEM graduates: some benchmark projections. Warwick: Department for Innovation, Universities and skills.

Wolff, P., Medin, D. L., & Pankratz, C. (1999). Evolution and devolution of folkbiological knowledge. *Cognition*, *73*(2), 177–204. doi:10.1016/S0010-0277(99)00051-7

Wood-Robinson, C. (1994). Young People's Ideas about Inheritance and Evolution. *Studies in Science Education*, *24*(1), 29–47. doi:10.1080/03057269408560038

Wright, I., Waterman, M., Prescott, H., & Murdoch-Eaton, D. (2003). A new Stroop-like measure of inhibitory function development: Typical developmental trends. *Journal of Child Psychology and Psychiatry*, *44*(4), 561–575.

Yeniad, N., Malda, A., Mesman, J., van IJzendoorn, M.H., & Piper, S. (2013). Shifting ability predicts maths and reading performance in children: a meta-analytical study. Learning and Individual Differences, 23(0), 1-9.

Younger, B.A., & Fearing, D.D. (1999). Parsing items into separate categories: developmental change in infant categorization. Child Development, 70(2), 291-303.

Yip, C.-W. (2009). Causal and teleological explanations in biology. *Journal of Biological Education*, *43*(4), 149–151. doi:10.1080/00219266.2009.9656174

Zaitchik, D., Iqbal, Y., & Carey, S. (2014). The Effect of Executive Function on Biological Reasoning in Young Children: An Individual Differences Study. *Child Development*, *85*(1), 160–175. doi:10.1111/cdev.12145

Zaitchik, D., Iqbal, Y., & Carey, S. (2013). The Effect of Executive Function on Biological Reasoning in Young Children: An Individual Differences Study. *Child Development*, *00*(0), 1–16. doi:10.1111/cdev.12145

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): a method of assessing executive function in children. *Nature Protocols*, *1*(1), 297–301. doi:10.1038/nprot.2006.46

Zheng, X., Swanson, H. L., & Marcoulides, G. A. (2011). Working memory components as predictors of children's mathematical word problem solving. *Journal of Experimental Child Psychology*, *110*(4), 481–98. doi:10.1016/j.jecp.2011.06.001

APPENDICES

A.1 – Examples of six contextual scenes used in Pilot Study 1.

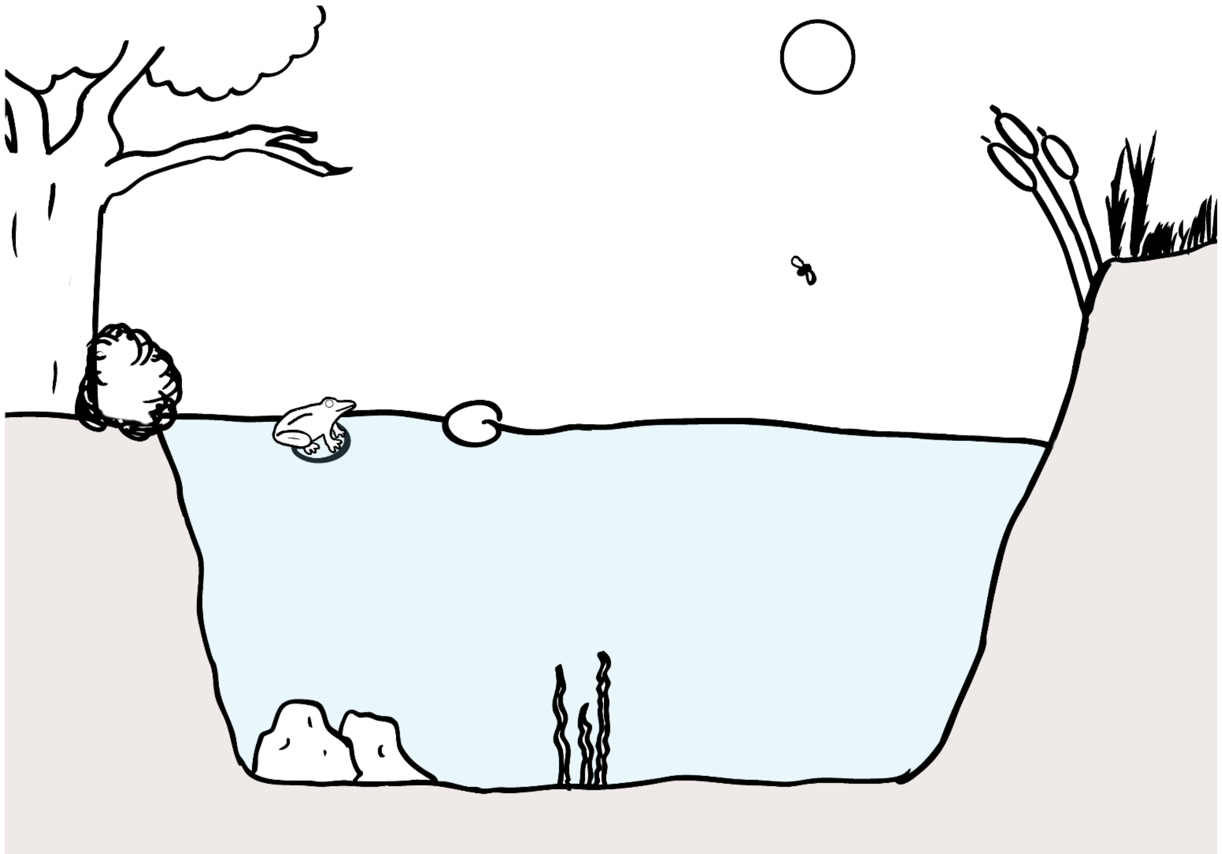


Figure A.1. Pond scene

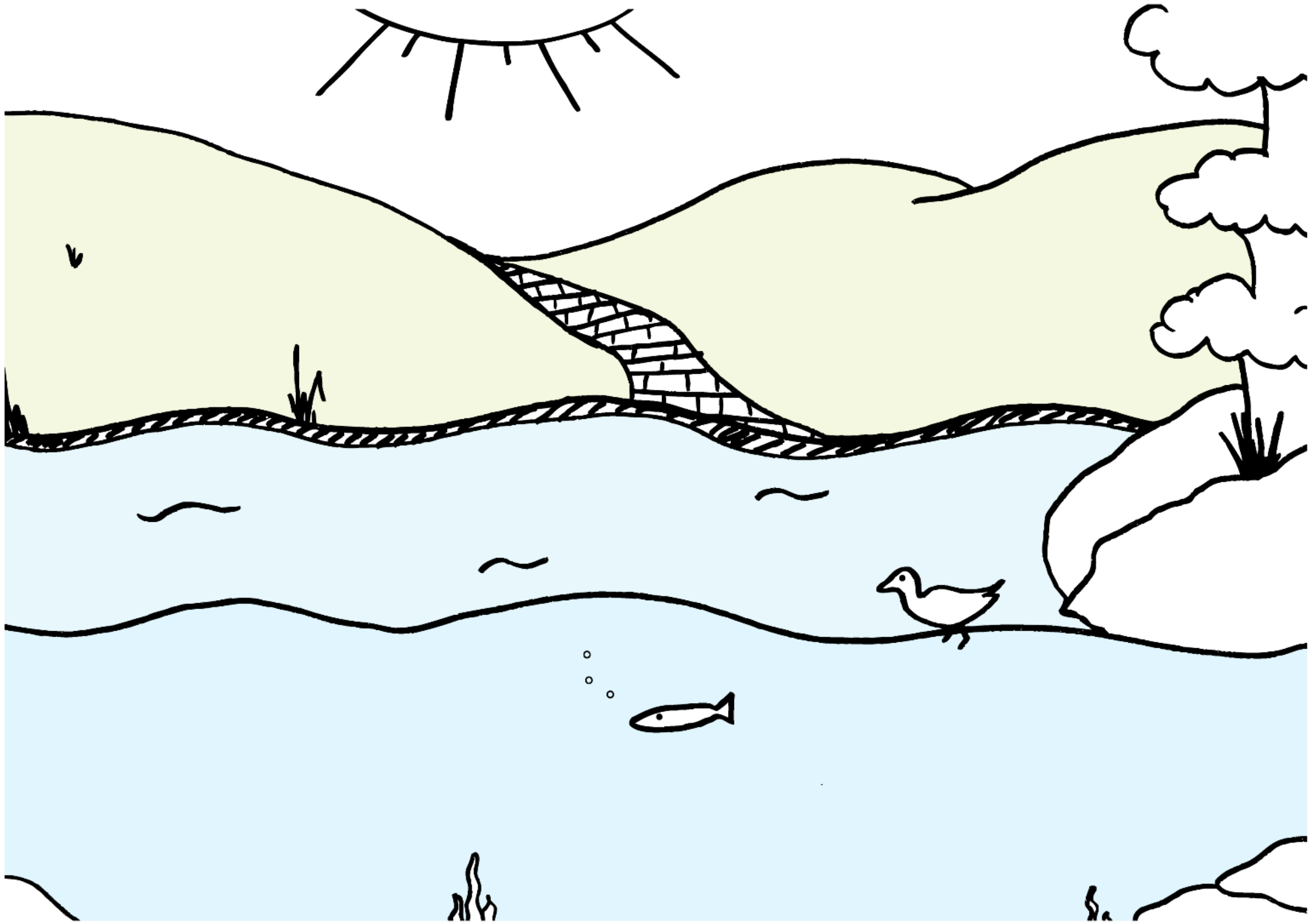


Figure A.2. Pond scene taken from Hipkins et al., 2008

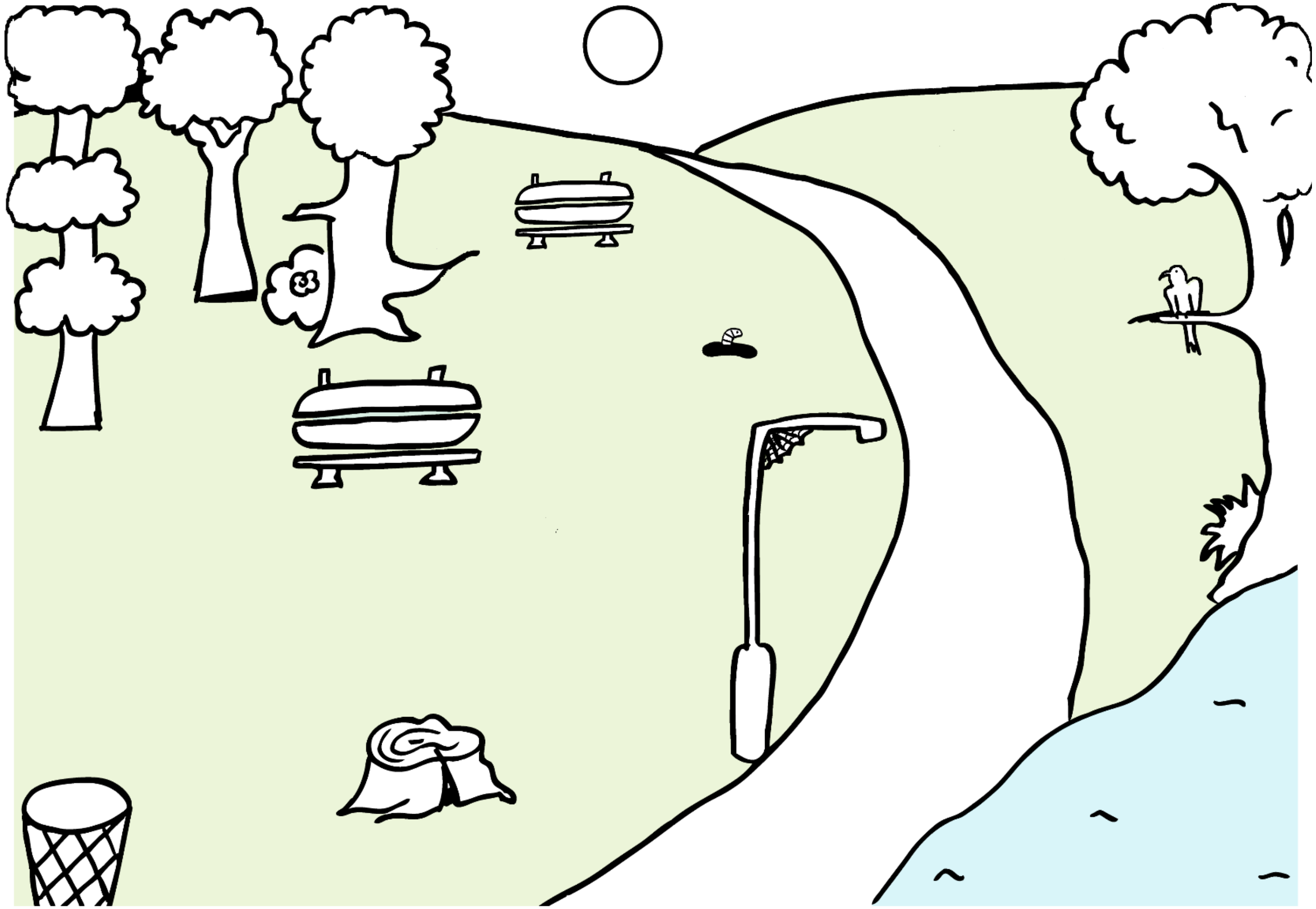


Figure A.3. Field Scene.

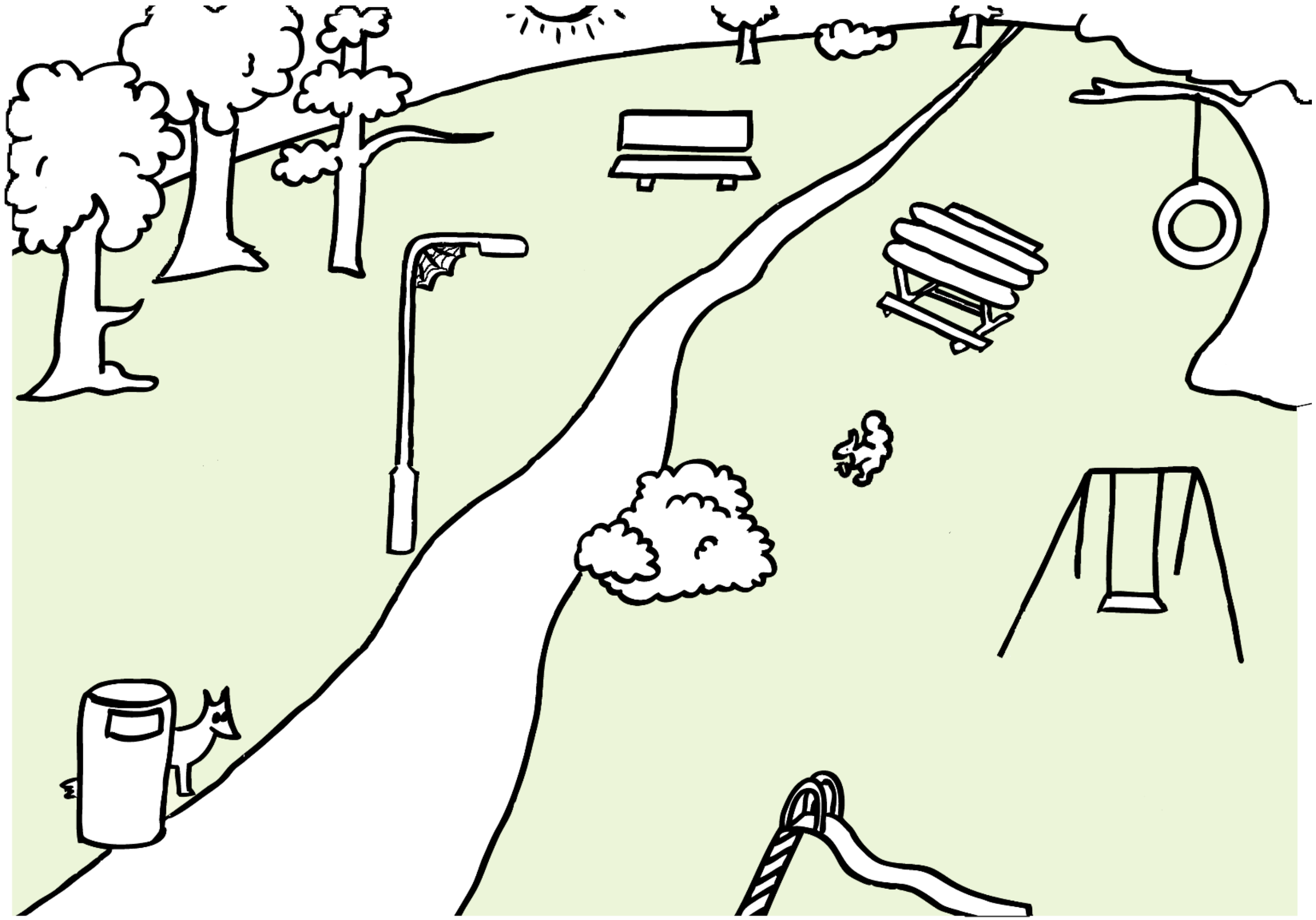


Figure A.4. Park scene

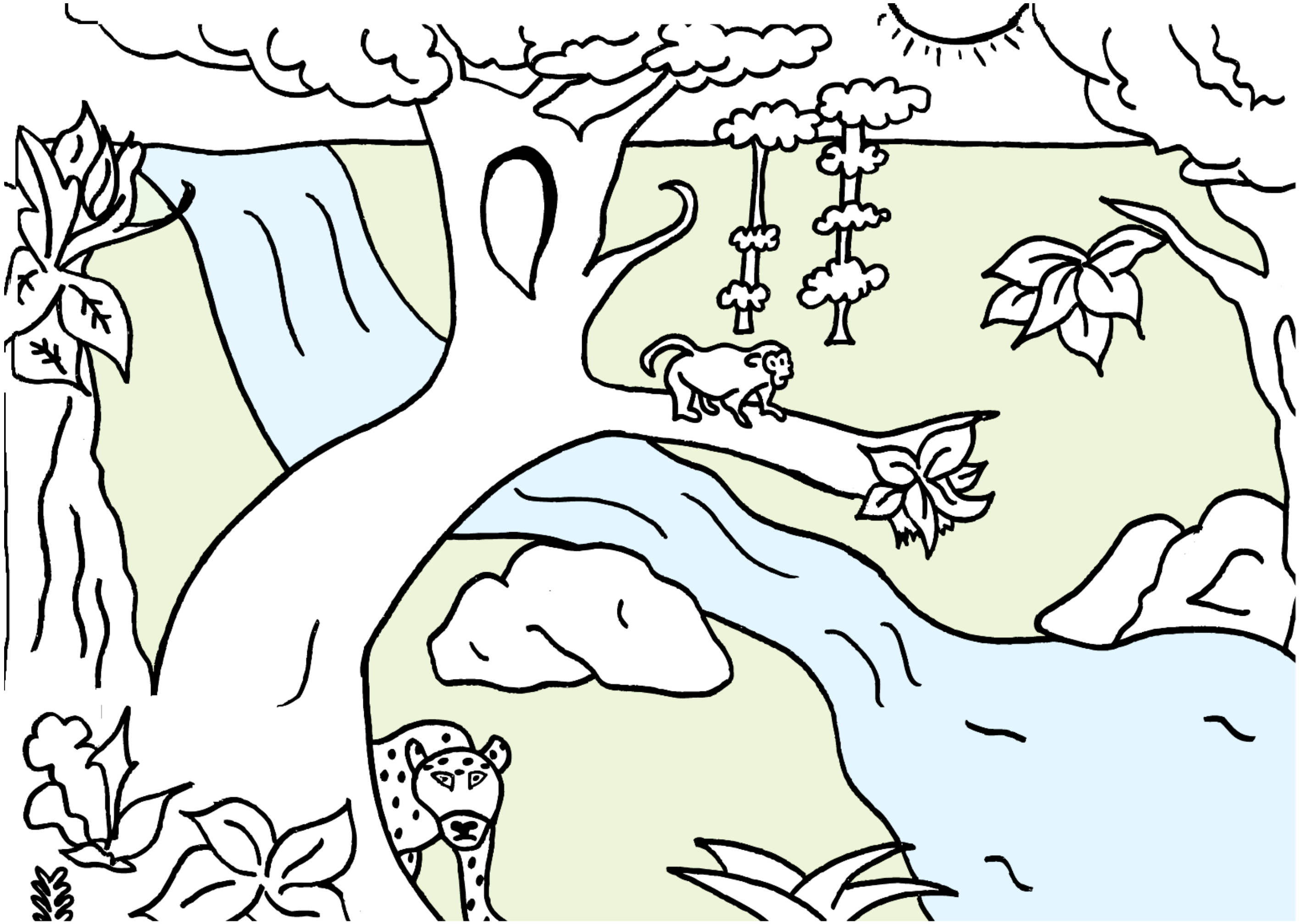


Figure A.5. Rainforest scene

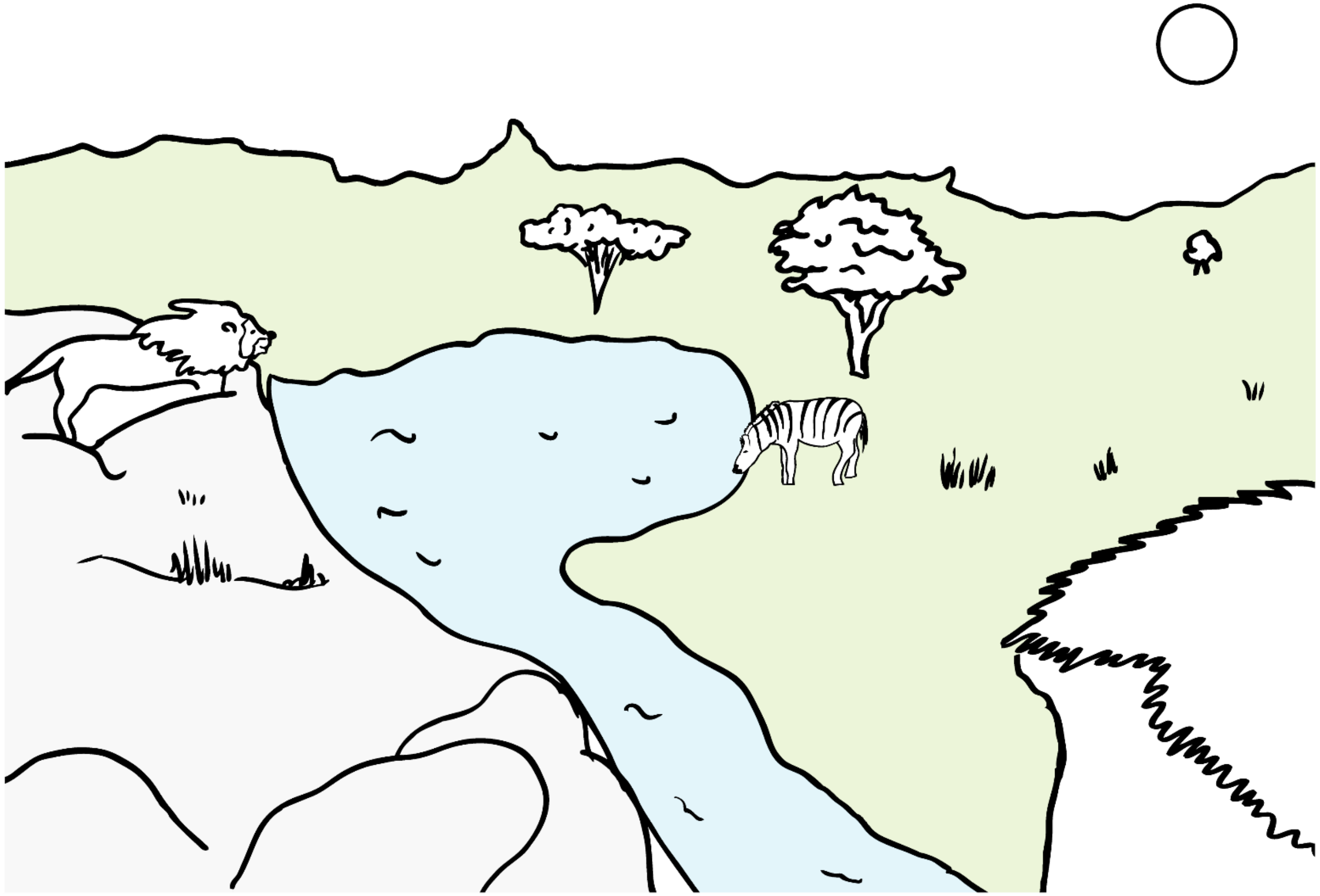


Figure A.6. Savannah scene

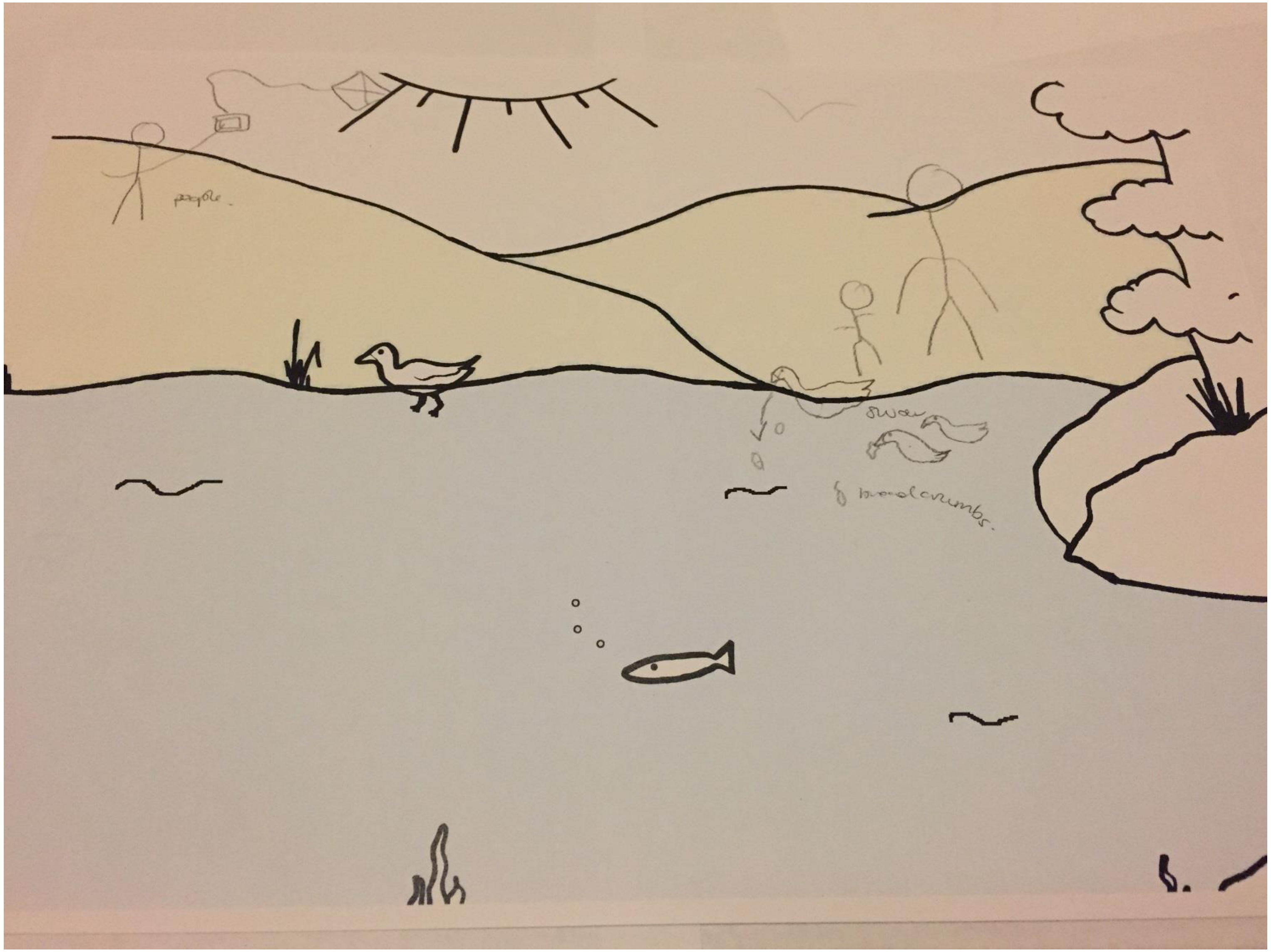
A.2 – Examples of children's drawings from Pilot Study 1.

Field context: focus is on drawing people

Savannah context: people were drawn alongside animals

Lake context: inaccurate animals included e.g. shark/crocodile.

Park context: heavily human-focused.



Lake context: very few animals

A.3 - Letters & questionnaires sent home to parents

Invitation to participate in research on primary school children's learning in biology

Dear parent/guardian,

Your child is invited to participate in a doctoral research project aimed at improving Science education in primary schools. The research is funded by the Economic and Social Research Council (ESRC), a government funded organisation.

The project will take a maximum of two hours, and simply involves giving children a number of literacy and numeracy based tasks and then an activity aimed at investigating children's knowledge of various biological phenomena. The children will be informally interviewed about their understanding of certain aspects of biology during this activity (e.g. biodiversity, ecology), and these will be audio-recorded. Any data your child provides will remain strictly anonymous and confidential which means that if the study is published, there will be no way of identifying any data as that of your child's. Please note that **this study is not in any way an assessment of your child's academic performance** and so no information can be provided about this. This is a longitudinal study and your child will be re-invited to participate in autumn term 2014, unless you specify otherwise.

Participation will be absolutely voluntary. If your child does not wish to answer questions or if they want to stop at any point, they are free to do so. If at any later point you wish your child's data to be withdrawn from the research, please let me know.

If you give permission for your child to participate in this longitudinal research, please complete the form below and return it to the school as soon as possible.

For further information you are welcome to contact either the head teacher of your child's school or myself at the Department for Psychology and Human Development, Institute of Education.

Yours sincerely,

Zayba Ghazali

(Doctoral researcher)

Professor Andy Tolmie

(Supervisor and Dean of the Doctoral School)

Email: zghazali@ioe.ac.uk

I _______________ (parent/guardian) give my full permission for

_________________________________ (full name) class ____________, to participate in your research.

Signed: ____________________________

Thank you.

Below is a short questionnaire to give us a little more information about your child, and the activities your child undertakes at home. **Any information you supply will remain strictly confidential and <u>anonymous</u>**. If at any point you wish to withdraw your responses, please let me know. If you have any questions or require further information, please contact me via email at *zghazali@ioe.ac.uk*

Thank you.

Please tell us about your household

1) How many adults are present in the home? _____

2) Please provide details of occupation and highest level of education of each adult in the home.

Adult one:	
Relationship to the child	
Occupation	
Highest level of education	
Adult two:	
Relationship to the child	
Occupation	
Highest level of education	
Adult three:	
Relationship to the child	
Occupation	
Highest level of education	
Adult four:	
Relationship to the child	
Occupation	
Highest level of education	

3) How many other children are present in the home? _____
Please provide details below

Child one	
Relationship to the child	
Age	
Child two	
Relationship to the child	
Age	
Child three	
Relationship to the child	
Age	
Child four	
Relationship to the child	
Age	
Child five	
Relationship to the child	
Age	

Please tell us about your child

4) Did your child attend pre-school? **Yes / No** (delete as appropriate)

If yes, please provide details by ticking the appropriate box below:

Private nursery	
State run nursery	
Play group	
Day care	
Other (please specify)	

5) Does your child receive free school meals? **Yes / No** (delete as appropriate)

6) Is English your child's first language? **Yes / No** (delete as appropriate)

   If no, please specify your child's first language:

   ____

7) What language is regularly spoken in your home? Please tick the appropriate box below:

English	
Other (please specify)	
A mixture (please specify)	

A.4 – Teacher Questionnaires

Year you teach: _______

Dear teachers,

Below is a short questionnaire about your experiences teaching science, and the recent changes in the primary science curriculum. **All responses are completely anonymous and confidential**. By completing and returning the questionnaire, you are consenting to participate in this research. Should you at any point wish to withdraw your responses or would like more information, please do not hesitate to contact me at: *zghazali@ioe.ac.uk*

What is your highest level of qualification in Science and Mathematics?	**Science**:______ **Maths**:_______
On average, approximately how many hours a week do you spend teaching science?	_____hours per week
On average, approximately how many hours a week do you spend teaching biology in particular?	_____hours per week

	Very Frequently	Frequently	Occasionally	Rarely	Never
In general, how often do you try to introduce novel scientific vocabulary when relevant to the lesson?					
On average, how often do you incorporate scientific vocabulary when teaching science in general?					
In general, how often do children experience "working scientifically" in science lessons?					
In general, how often do children experience "working scientifically"					

during biology lessons specifically?					

By the end of the school year, approximately what percentage of time teaching science will you have spent during this school year on each of the following content areas? (The total should add to 100%)	
Biology (e.g. structure/function; life processes, reproduction/heredity, natural selection; ecosystems, human health)	%
Chemistry (e.g. classification, composition and properties of matter; chemical change)	%
Physics (e.g. physical states/ changes in matter; energy; light; sound; electricity and magnetism; forces and motion)	%
Other, please specify	%

How often do you incorporate ICT/video materials in science class to do the following:

	Very Frequently	Frequently	Occasionally	Rarely	Never
Drill and practice					
Demonstrate scientific principles					
Play science learning games					
Do laboratory simulations					
Take a test or quiz					
In general what methods do you employ when teaching scientific topics and how often do you use these techniques?					
Have children observe natural phenomena and describe/record information about what they see					

Let children watch you demonstrate an experiment/investigation					
Have children design or plan experiments or investigations					
Have children conduct experiments or investigations					
Work together in small groups (of mixed ability) on experiments or investigations to come up with a joint solution to a problem					
Work together in small groups (based on ability) on experiments or investigations to come up with a joint solution to a problem					
Read their textbooks or other resource material					
Have students memorise facts and principles					
Have children give explanations about something they are studying					
Relate what they are learning in science to their daily lives					

How often do you do each of the following activities during a science lesson?

	Very Frequently	Frequently	Occasionally	Rarel y	Never
Conduct pre-assessment to determine what students already know					

Present new topics to a the class lecture-style					
Explicitly state learning goals					
Give different work to the students that have difficulties learning and/or to those who can advance faster					
At the beginning of the lesson, present a short summary of the previous lesson					
Work with individual students					
Get students to hold a debate and argue for a particular point of view which may not be their own					
Assess learning by holding a test or quiz					
Review students' homework *with the students*					

In the past three years, have you participated in professional development in any of the following?

	Yes	No
Science content		
Science pedagogy and instruction		
Science curriculum		
Integrating ICT into science		
Improving students' critical thinking or inquiry skills		
Science assessment		

How would you rate the impact of any training/CPD in science education that you have received in each of the following areas?

	Not	Confirmed what I	Caused me to change my

	applicable	was already doing	teaching practices
Deepening my own science content and knowledge			
Understanding student thinking in science			
Learning how to use enquiry/investigation-orientated teaching strategies			
Learning how to use technology in science instruction			
Learning how to assess student learning in science			
Learning how to teach science in a class that includes students with special needs			

How well qualified do you feel to teach each of the following subjects at the Key Stage you teach, whether or not they are currently included in your curriculum?

	Very well qualified	Adequately qualified	Not as well qualified, and would like more training
Biology			
Chemistry			
Physics			
Mathematics			
Literacy/reading			
Social sciences			

Have you noticed any significant differences in the new science curriculum that have caused you to alter/modify your past teaching methods? Please describe.

In general how easy have you found introducing scientific topics from the new curriculum into the classroom?
Finally, do you have any other comments regarding the changes in the science curriculum, your experiences, or observations of children's consequent learning in the classroom?

Thank you for completing this questionnaire and for all your patience and generosity throughout the testing period. If you would like to find out more about the project and the findings published from the longitudinal study, please email me at: *zghazali@ioe.ac.uk*

Zayba Ghazali
(doctoral researcher)

A.5 – Qualitative analysis of the teacher data

As discussed at the end of Chapter 7, the return rate of completed teacher questionnaires was very low, with only 13 out of 65 questionnaires having been returned, a response rate of only 20% (see Table 6.1 in Chapter 7). Aside from the Reception year group where no teachers returned the questionnaire, at least one questionnaire was returned for each academic year group across the three schools. For these reasons, data were explored qualitatively and interpreted with caution. The data was also considered by year group and not by school because of the poor sample size.

1 Qualifications

There were a total of three questionnaires returned from Year 1 teachers. Of these the highest level of qualification for science was at A-Level, whereas for maths, one teacher was qualified up to degree level. One questionnaire was returned from a Year 2 teacher, who was qualified to GCSE level for both maths and science, as were all three Year 3 teachers. Year 4, Year 5, and Year 6 both had at least one teacher qualified at A-Level for either maths or science. Generally however, it appears as though GCSE level seems to be the highest level of qualification across the majority of primary school teachers in this sample.

2 Teaching hours

Nearly all teachers reported spending approximately 1-2 hours teaching science per week. For biology in particular, 30 minutes to 1 hour per week of teaching time was spent, although many teachers reported this was topic dependent. Generally all teachers considered there to be equal teaching time for biology, chemistry, and physics topics (approximately 33.3%), however some teachers (particularly those in Years 2 and 5) considered a larger proportion of teaching time to be placed on biology topics specifically, approximately 50-70% of science teaching. This may be as a result of specific topics that are introduced in those particular year groups that might require extra attention.

3 Information and Communication Technology (ICT) and science

In the questionnaires, teachers were asked how often they employed the use of ICT in their science lessons. They were asked to report the frequency of techniques including: drill and practice, virtual demonstrations, science games, laboratory simulations, and tests/quizzes. The majority of teachers across all years appeared to report similarly with regards to their ICT use in science lessons. For drill and practice, very few teachers reported using this technique. By far the majority of teachers seemed to use ICT for providing virtual demonstrations and allowing children to play science related games. ICT was rarely used for laboratory simulations or testing across all year groups.

4 Teaching strategies

Reporting on the teaching strategies used in science lessons was the main focus of the questionnaire. Teachers were asked to rate how frequently they used certain strategies or techniques. These will be explored separately for each technique.

Introduce relevant novel scientific vocabulary:

All teachers across all year groups reported introducing new and relevant scientific vocabulary in lessons *very frequently* or *frequently*. The same was also true in biology lessons in particular. This may be a consequence of some of the curricular changes introduced in 2014 that emphasise the use of vocabulary and scientific terms.

How often children experience working scientifically:

In terms of how often children experienced working scientifically, the majority of teachers consistently reported *frequently* or *occasionally*. With regards to working scientifically in biology lessons, most teachers reported this was *occasionally*, with teachers in Year 4 and Year 6 reporting this was *rare*.

How often children observe natural phenomena & record it:

Reporting on use of this teaching strategy was slightly more variable with responses ranging from very *frequently* to *occasionally* both within and across year groups. However, it seems as though generally this is a technique most teachers employ.

How often children watch the teacher demonstrate an experiment:

Again, responses to this technique ranged from *frequently* to *rare*. Only the Year 2 teacher reported using this strategy rarely and as this was the only Year 2 teacher who responded to the questionnaire, it cannot be said for certain that this is typical of all teachers in Year 2. The majority of other teachers from all year groups seems to demonstrate experiments or investigations *frequently* and *occasionally* suggesting this is a popular technique.

How often children conduct experiments:

Similarly to watching teachers demonstrate experiments, the responses for how often children participated in experimental investigation themselves was variable from very *frequently* to *rare*. There was only one teacher in Year 1 who reported this as a *rare* practice, whilst the other Year 1 teachers reported occasional use. It might be that Year 1 children are younger and so cannot conduct many experimental investigations safely or indeed appropriately. The frequency for use of this technique seemed to increase with the increasing age of the child, with teachers in Year 6 reporting *very frequent* use.

How often children work in mixed ability or ability groups on experiments to come up with a joint solution:

The responses for how often teachers made children work in either mixed ability groups or ability groups appeared to be the same, hence the results are collapsed under one heading. Generally this seemed to be a popular strategy with teachers across all years reporting *very frequently* to *occasionally*. There were slightly more *occasionally* responses for splitting children into ability groups than mixed ability groups suggesting teachers preferred mixed ability groupings.

How often children read textbooks/resources and memorise facts/principles:

The use of textbooks or other resource material was reported as *occasionally* to *never* across all year groups. It may be that individual learning from resource material is typical of children in secondary school, rather than in primary school where it seems as though teachers actively guide the learning of children, at least within science lessons. Likewise, teachers reported making children memorise facts and principles very *occasionally* to *never* perhaps for the same reasons.

How often children give explanations about something they are studying:

Getting children to provide explanations about something they are studying seemed to be a teaching strategy that was popular among the sample, with teachers reporting use as *very frequently* to *occasionally* across all year groups. It may that this is a tool used by teachers as a form of assessment, hence its frequent use.

How often children relate what they are learning in science to their everyday lives:

Relating science to children's everyday lives seemed to be less of a focus among teachers with only three from the whole sample reporting this as *frequently*. One teacher in Year 1 reported this as *rare*, whilst the majority reported *occasionally*. It may be that in many cases it is fairly obvious to children how certain things they learn relate to their everyday lives and so teachers may place less of a focus on this aspect.

How often topics are presented lecture-style:

Presenting new topics lecture style was not a popular teaching technique among the sample with the majority of reponses being *rare* to *never*, with very few teachers responding *occasionally*. It may be that this technique is not very effective with younger children who may require more stimulation and various modes of delivery in order to learn and remain attentive.

Learning goals are explicitly stated & a brief summary of the lesson before is provided:

Explicitly stating learning goals is something teachers across year groups seemed to do very often with all responses ranging from *very frequently* to *occasionally*. Indeed this is something that is recommended in the National Curriculum (NC; 2014) as good practice in order to make children aware of what they about to be/being taught. Likewise, providing a brief summary of the material taught in the previous lesson was popular with teachers' responses being either *very frequently* or *frequently*. Again, the NC (2014) considers this good practice, which would explain its popularity.

How often different work is given to students with learning difficulties/children who advance faster:

Providing different work to children with learning difficulties or extra work to children who advance faster is also something teachers do *very frequently* or *frequently* in this sample. It may be that this is something that is required in order to make sure children are able to work and learn effectively within a classroom environment.

Work individually with students:

Working individually with teachers was also something that teachers did *very frequently* to *occasionally* across all year groups. It is likely that this is a way of assessing children and being able to target those who are in need of more help.

Get children to hold debates and argue a viewpoint that might not be their own:

Responses to using this technique varied from *very frequently* to *never*. There seemed to be group differences with regards to this technique in that teachers of Year 1 and Year 2 reported very little to no use, whereas teachers from Year 3 and above reported frequent use. It may be that younger children are not able to articulate themselves well enough at this stage to debate, particularly one that is not their own. With increasing age, however, teachers recognise this as a beneficial technique.

Assess learning by holding a quiz:

Assessing learning through quizzes was reported as less frequent with responses ranging from *occasionally* to *never*. Teachers in Year 1 and Year 2 generally reported never using this strategy at all, where as those is Years 3 and above reported rare to occasional use. It could be that in the lower year groups children are not well managed in test conditions, or may not have sufficient knowledge to adequately test. With the older children this is still not a popular strategy, however, and it could be that teachers do not wish to cause unnecessary distress.

Review homework with the student:

Teachers responses for this were variable because children in the lower years often did not receive any homework, hence responses ranged from *occasionally* to *never.* It could be that teachers often do not have the time to sit with children individually and go through homework as a general teaching strategy and instead conduct informal assessments using other methods such as working individually with students and providing different work to those who had difficulties or advanced faster.

5 Professional training in science education

Nearly all teachers had undergone some professional training in science education specifically, apart from one teacher in Year 4 who had not yet received any. The areas which teachers had professional training seemed to vary both across and within year groups. This could reflect the topics in which the teachers felt they lacked knowledge and therefore sought training. The areas of training included: science content, pedagogy and instruction, science curriculum, integrating ICT into science lessons, improving children's critical thinking or inquiry skills, and science assessment. Although the aspects of professional training varied across teachers, nearly all teachers felt that training helped them to change their teaching practices for the better. Only few of the teachers felt that the training confirmed what they were already doing and this was only for some, not all areas, they received training in.

6 Teacher comments about the changes in the curriculum

Of the thirteen teachers that returned their questionnaires, only five provided any comment about their views on the curricular changes regarding primary science. These were two teachers from Year 4, two from Year 5, and one from Year 1.

The Year 1 teacher felt that the new curriculum seemed more prescriptive and somewhat restrictive in nature, especially in terms of the new vocabulary to be taught. This teacher also felt that for Year 1 in particular, all new topics were easy to introduce mainly because the science coordinator at the school had provided abundant resources in preparation for the new changes.

The year 4 teachers noted that the changes allowed more group work and pupil exploration, which they felt was a benefit. They also noted a focus on inquiry skills and critical thinking, yet with regards to how easy introducing new topics were, one teacher felt they lacked resources at times.

In contrast, the Year 6 teachers felt that the new curriculum lacked objectives, was more knowledge based, and therefore limited the opportunity for children to access science inquiry skills. They note however, that the school in general promotes this as a focus and have provided training to aid teachers in this endeavour. One teacher went on to write, "*the new science curriculum has low expectations of primary school children*...[there needs to be] *more chance for experimentation/working scientifically*."

In sum it appears though teachers are generally of mixed opinion about the new changes. This, again, may be due to the year-specific topics that teachers have had to introduce, and to some extent, the support, training, and resources they have received from their school.

A.6 - Results from Time 1 regression models with new parent composite variable (10.6.1.1)

Table A.1. Hierarchical regression model for biodiversity at Time 1 with new parent composite variable

Biodiversity Time 1		
Model 1: AdjR^2=0.542 (*p*<0.001)		
	Beta	sig
Age	0.715	**<0.001**
No. of adults	0.049	0.518
No. of younger children	0.078	0.286
English native	-0.168	**0.030**
Parent edu/occ	-0.012	0.875
Model 2: AdjR^2=0.717 (*p*<0.001)		
Age	0.446	**<0.001**
No. of adults	0.018	0.759
No. of younger children	0.036	0.539
English native	-0.041	0.521
Parent edu/occ	-0.050	0.402
Ecology	0.421	**<0.001**
Inheritance	-0.058	0.453
Evolution	0.180	0.069

Table A.2. Hierarchical regression model for ecology at Time 1 with new parent composite variable

Ecology at Time 1		
Model 1: AdjR^2=0.244 (*p*<0.001)		
	Beta	sig
Age	0.454	**<0.001**
No. of adults	0.065	0.501
No. of younger children	0.062	0.508

English native	-0.206	**0.038**
Parent edu/occ	0.06	0.539

Model 2: AdjR^2=0.651 (*p*<0.001)		
Age	-0.216	**0.028**
No. of adults	0.048	0.470
No. of younger children	-0.024	0.715
English native	0.017	0.807
Parent edu/occ	0.010	0.885
Biodiversity	0.518	**<0.001**
Inheritance	0.295	**<0.001**
Evolution	0.319	**0.003**

Table A.3. Hierarchical regression model for inheritance at Time 1 with new parent composite variable

Inheritance at Time 1		
Model 1: AdjR^2=0.200 (*p*<0.001)		
	Beta	sig
Age	0.407	**<0.001**
No. of adults	-0.033	0.739
No. of younger children	0.042	0.667
English native	-0.163	0.108
Parent edu/occ	0.081	0.416
Model 2: AdjR^2 = 0.396 (*p*<0.001)		
Age	0.178	0.175
No. of adults	-0.061	0.481
No. of younger children	0.004	0.965
English native	-0.037	0.691
Parent edu/occ	0.034	0.698
Biodiversity	-0.123	0.453

Ecology	0.51	**<0.001**
Evolution	0.151	0.298

Table A.4. Hierarchical regression model for evolution at Time 1 with new parent composite variable

Evolution Time 1		
Model 1: AdjR^2=0.464 (*p*<0.001)		
	Beta	sig
Age	0.565	**<0.001**
No. of adults	0.006	0.942
No. of younger children	0.104	0.192
English native	-0.276	**0.001**
Parent edu/occ	0.101	0.219
Model 2: AdjR^2=0.626 (*p*<0.001)		
Age	0.219	**0.028**
No. of adults	-0.023	0.727
No. of younger children	0.062	0.341
English native	-0.156	**0.027**
Parent edu/occ	0.077	0.254
Biodiversity	0.226	0.069
Ecology	0.326	**0.003**
Inheritance	0.089	0.298

A.7 - Results from Time 2 regression models with new parent composite variable (10.6.1.2)

Table A.5. Hierarchical regression model for biodiversity at Time 2 with new parent composite variable

Biodiversity Time 1		
Model 1: AdjR^2=0.577 (*p*<0.001)		
	Beta	sig
Age	0.649	**<0.001**
No. of adults	-0.130	0.081
No. of children	0.008	0.911
English native	-0.186	**0.017**
Parent edu/occ	0.149	**0.048**
Model 2: AdjR^2=0.780 (*p*<0.001)		
Age	0.194	**0.012**
No. of adults	-0.170	**0.002**
No. of younger children	0.082	0.138
English native	-0.037	0.522
Parent edu/occ	0.076	0.182
Ecology	0.388	**<0.001**
Inheritance	0.021	0.776
Evolution	0.317	**0.001**

Table A.6. Hierarchical regression model for ecology at Time 2 with new parent composite variable

Ecology at Time 1		
Model 1: AdjR^2=0.508 (*p*<0.001)		
	Beta	sig
Age	0.663	**<0.001**
No. of adults	0.041	0.610
No. of younger children	-0.111	0.163

English native	-0.188	**0.024**
Parent edu/occ	0.120	0.138

Model 2: AdjR^2=0.713 (p<0.001)		
Age	0.201	**0.024**
No. of adults	0.084	0.196
No. of younger children	-0.103	0.099
English native	-0.045	0.498
Parent edu/occ	0.042	0.525
Biodiversity	0.505	**<0.001**
Inheritance	-0.060	0.470
Evolution	0.266	**0.017**

Table A.7. Hierarchical regression model for inheritance at Time 3 with new parent composite variable

Inheritance at Time 1		
Model 1: AdjR^2=0.301 (p<0.001)		
	Beta	sig
Age	0.411	**<0.001**
No. of adults	-0.027	0.776
No. of younger children	-0.176	0.065
English native	-0.173	0.081
Parent edu/occ	0.247	**0.011**
Model 2: AdjR^2 = 0.441 (ps <0.001)		
Age	0.126	0.314
No. of adults	-0.058	0.527
No. of younger children	-0.143	0.101
English native	-0.061	0.510
Parent edu/occ	0.216	**0.016**
Biodiversity	0.052	0.776

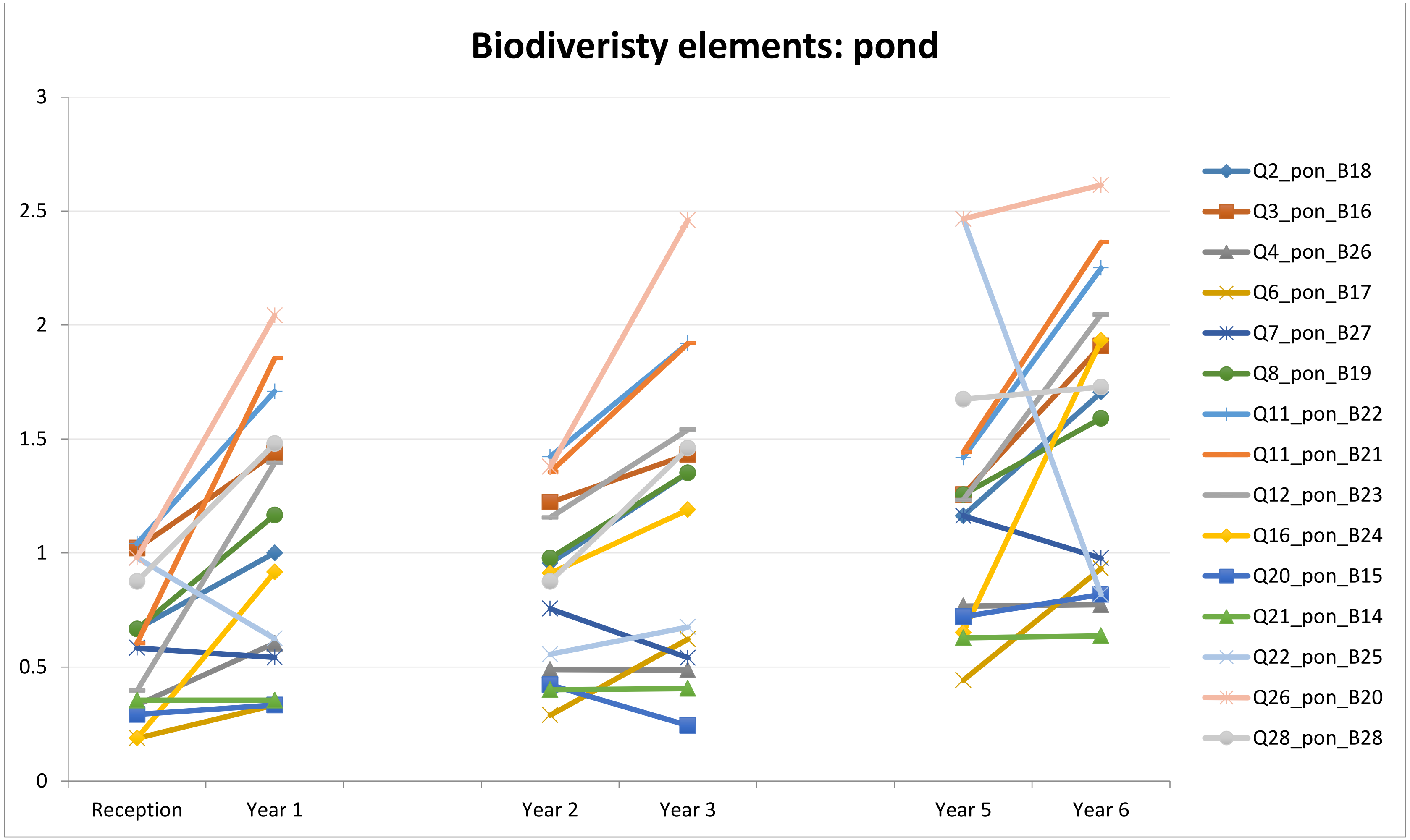
Ecology	-0.117	0.470
Evolution	0.548	**<0.001**

Table A.8. Hierarchical regression model for evolution at Time 2 with new parent composite variable

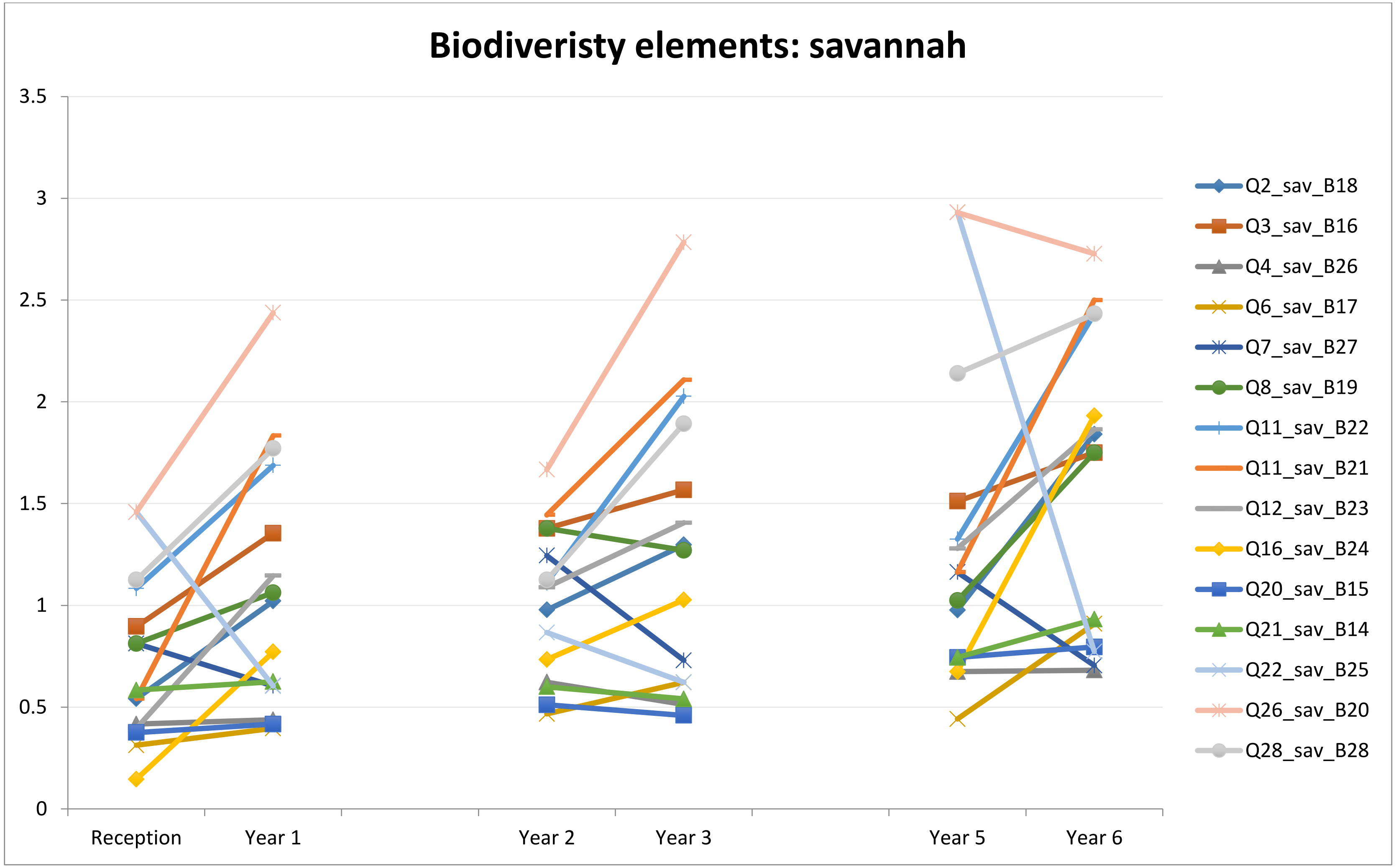
Evolution Time 1		
Model 1: AdjR^2=0.409 (*p*<0.001)		
	Beta	sig
Age	0.598	**<0.001**
No. of adults	0.077	0.379
No. of younger children	-0.084	0.333
English native	-0.227	**0.014**
Parent edu/occ	0.067	0.445
Model 2: AdjR^2=0.701 (*p*<0.001)		
Age	0.015	0.871
No. of adults	0.129	**0.049**
No. of younger children	-0.005	0.935
English native	-0.044	0.515
Parent edu/occ	-0.102	0.123
Biodiversity	0.43	**0.001**
Ecology	0.277	**0.017**
Inheritance	0.293	**<0.001**

A.8 – Elements graphs

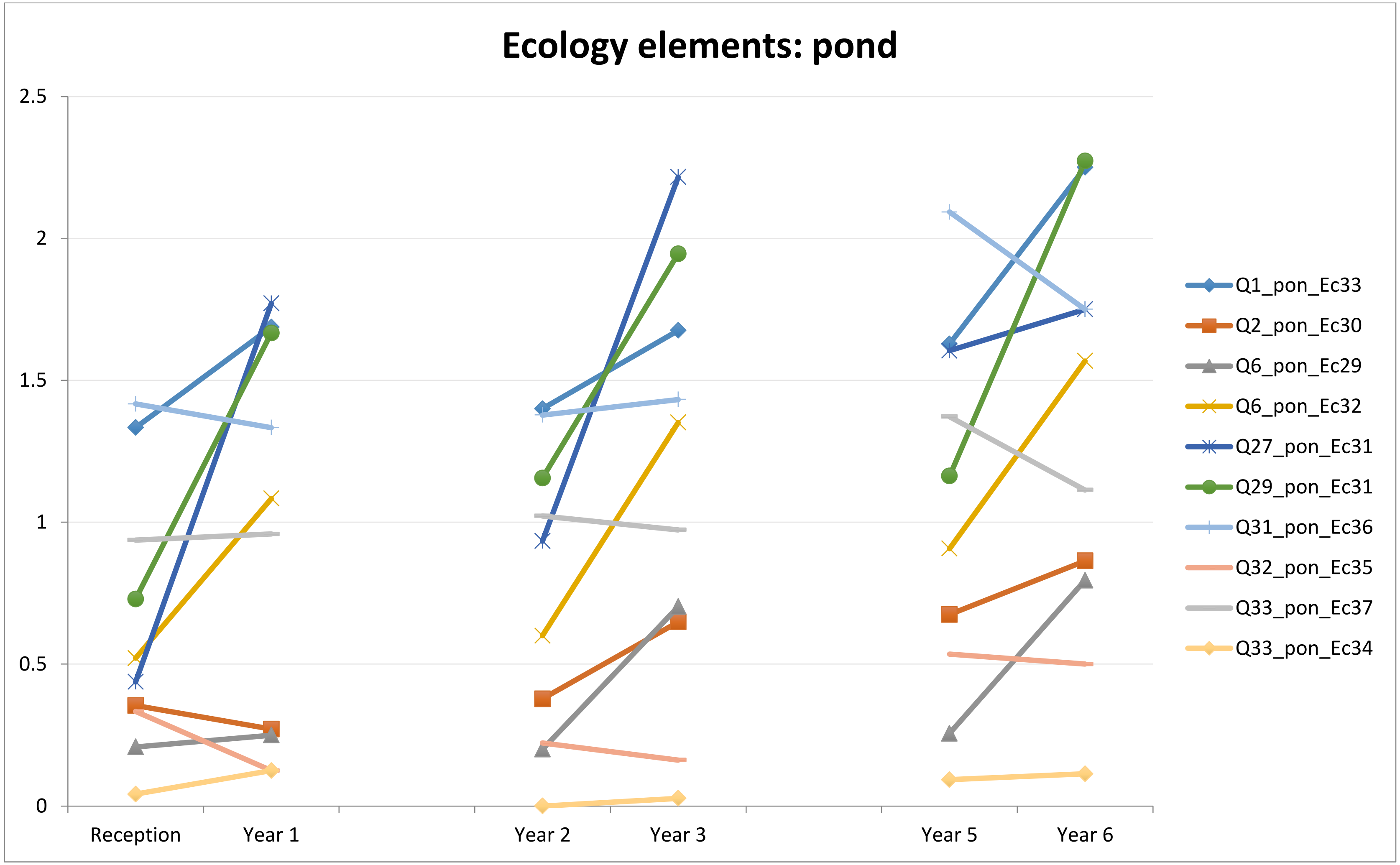
Biodiversity pond context



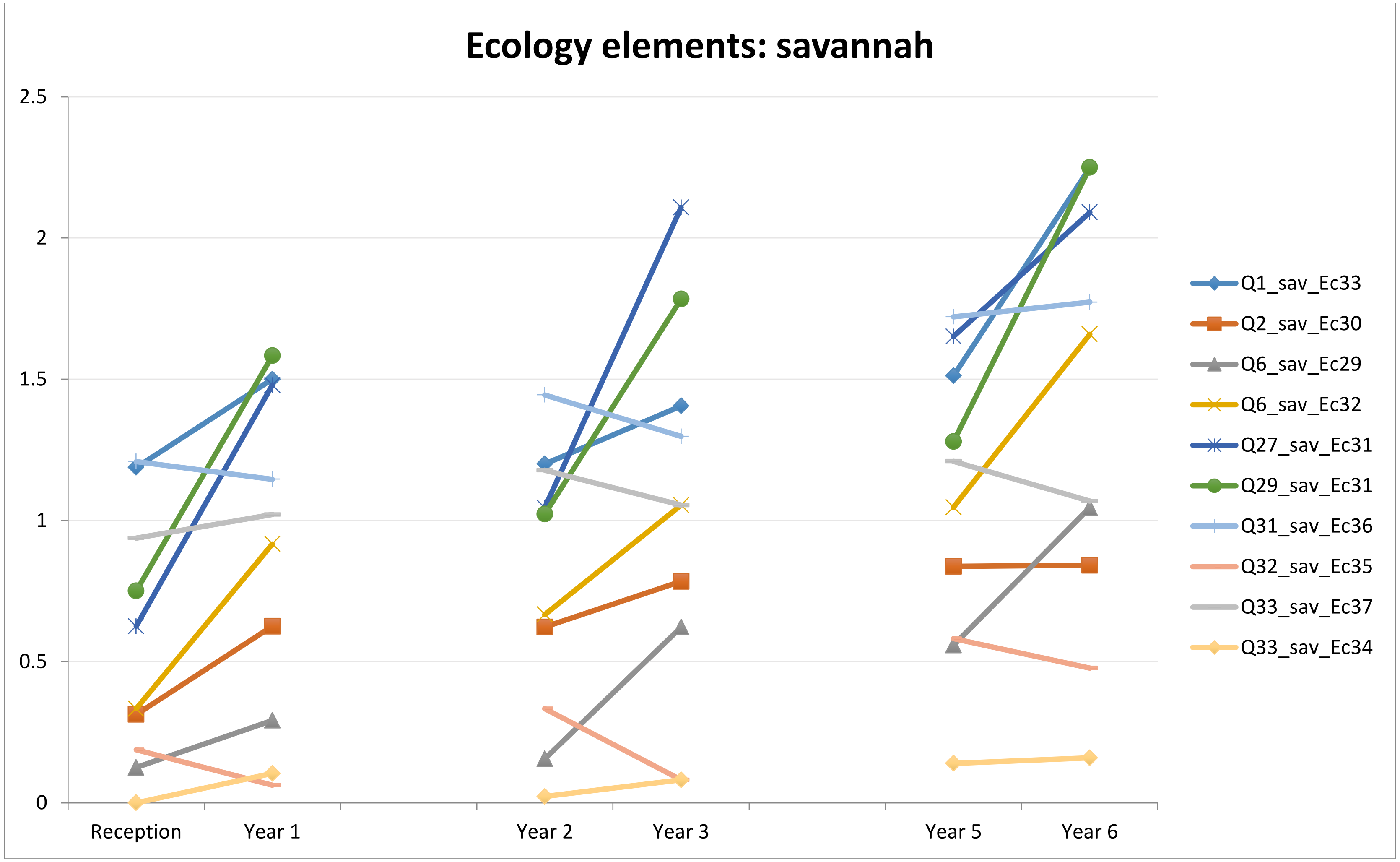
Biodiversity savannah context



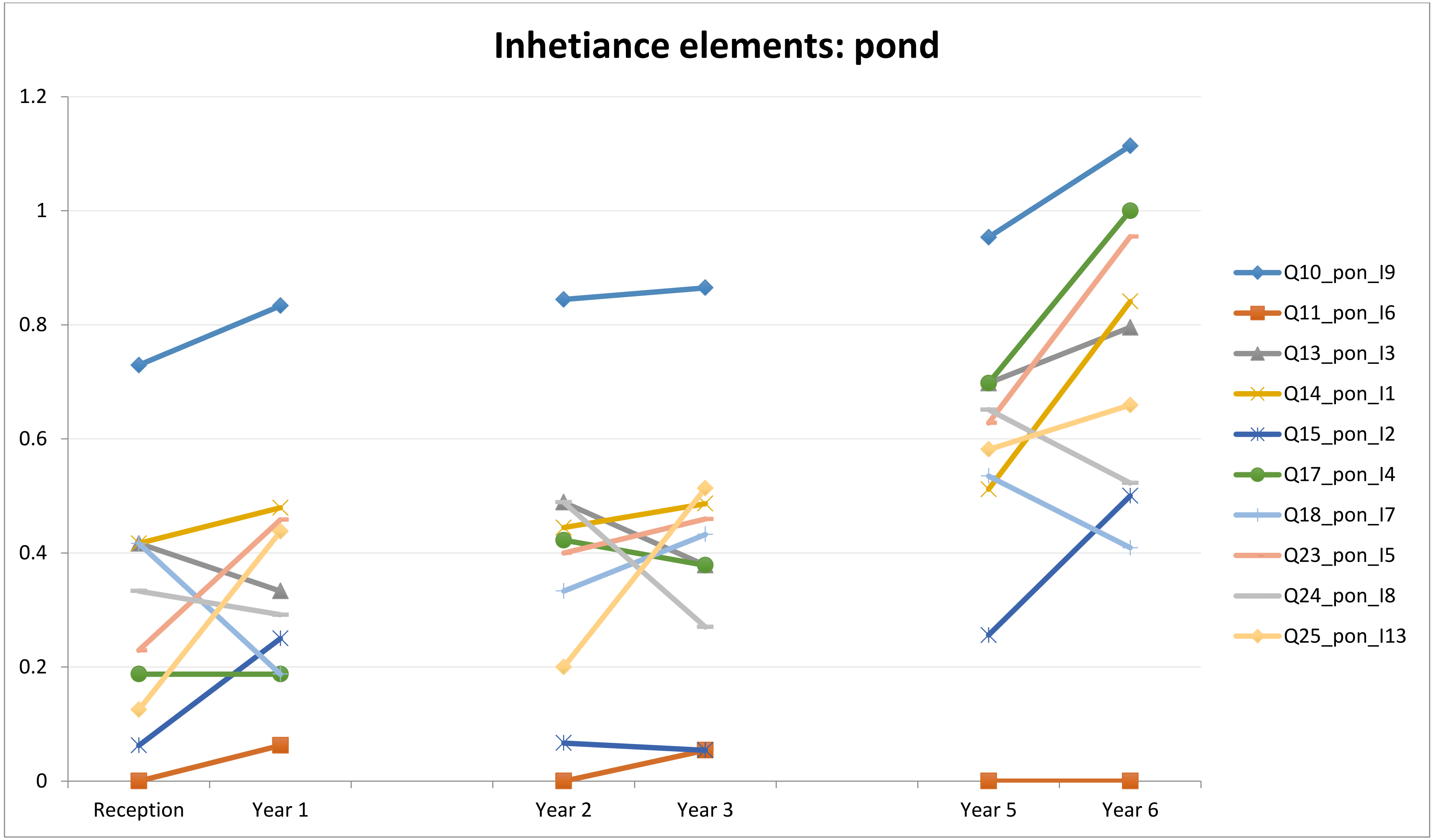
Ecology pond context:



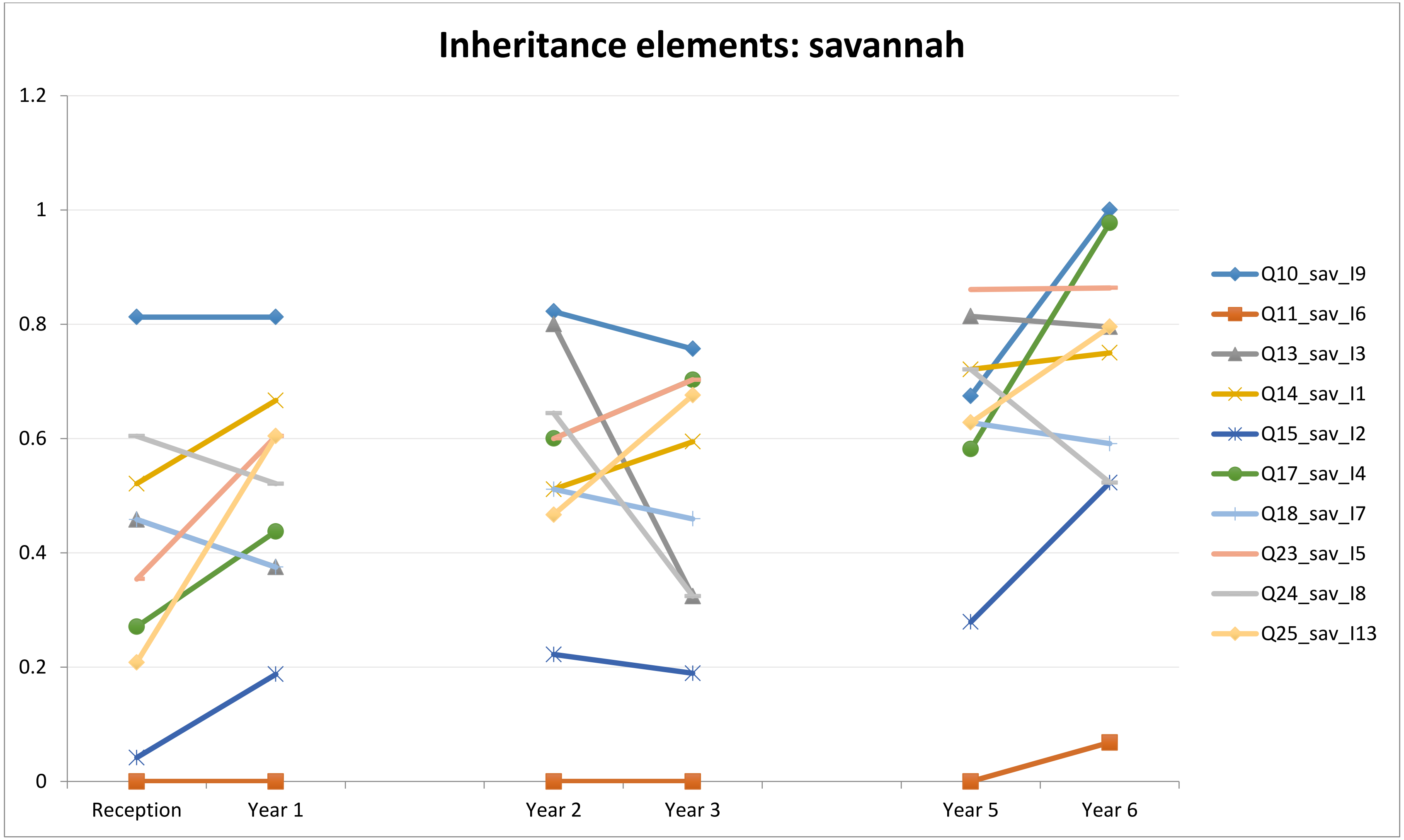
Ecology savannah context:



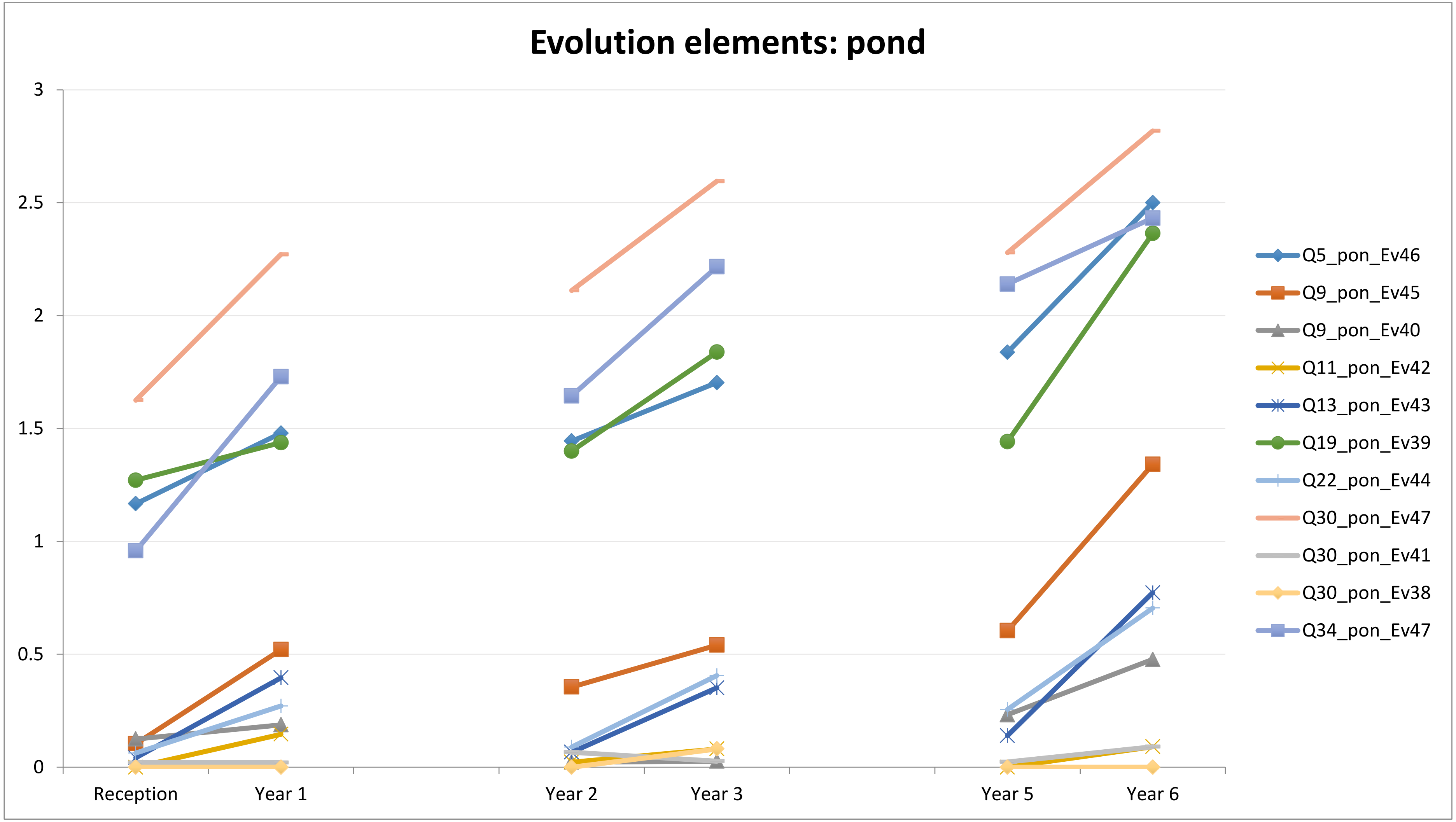
Inheritance pond context:



Inheritance savannah context:



Evolution pond context:



Evolution savannah context: